The Distributional Effects of Illness and Air Pollution

Borghild Marie Moland Gaarder

Ph.D. Thesis in Economics

University College London

2002
Acknowledgements

Four years have gone by while I was working on this thesis now in front of you. It comprises some eighty-two thousand words,¹ some of them describing new and hopefully worthwhile insights and ideas. A number of people have given me support, advice, and feedback during the process. I would like to thank my supervisors David Ulph and David Pearce at University College London for their guidance. More support was provided by Robert Mabro at the Oxford Institute for Energy Studies, Oxford, Torstein Bye, Knut-Einar Rosendahl, and Bente Halvorsen at Statistics Norway, Oslo, Gunnar Eskeland, Maureen Cropper, Mead Over, Andrew Sunil Rajkumar, Kirk Hamilton, and John Dixon at the World Bank, Washington DC, Frank Windmeijer at the Institute for Fiscal Studies (IFS), London, and Ulrich Bartsch throughout the process. My special thanks to my parents, Tora Moland Gaarder and Jon A. Gaarder, for always being there and readily providing emotional and intellectual support.

The meta-analysis described in the last part (chapter 6) of this thesis builds upon my work with David Maddison, Senior Research Fellow at the Centre for Social and Economic Research on the Global Environment (CSERGE), and his contribution is gratefully acknowledged. Chapter 6 is however not a co-authored paper – data sample, as well as methodology and variables differ from the study we carried out jointly. In addition to the above differences, the present paper reviews the existing meta-analysis literature, carries out a thorough sensitivity analysis, and discusses in detail the weaknesses attached to such a study. The relation between the present study and the co-authored study (Maddison and Gaarder, 2001) is described in detail in Appendix 6.A.


¹ Ninety-nine thousand words when counting references and appendices.
Abstract

This thesis explores the possible interactions of income and illness through labour supply and averting and mitigating behavioural choices, and the implications for health and health-cost measurements. In the first part, standard utility maximisation theory is used to analyse labour supply behaviour under constraints imposed by sickness and minimum consumption requirements. For a rather general utility function, an empirically supported elasticity of substitution between leisure and consumption, and under the assumption that no sick pay is received, we find that only higher-wage individuals will choose to recuperate fully, whereas others will work while being sick. Therefore, the income and welfare losses due to a given illness may be larger for low-wage individuals than for those with higher wages.

The second part discusses the important role that income may play in the quantification and valuation of the health effects from air pollution. The level of income may affect the amount of averting and mitigating activities undertaken, and the willingness to pay to avoid an adverse health effect. The validity of the methods used to quantify and value the health impact of air pollution is therefore questioned. Furthermore, the effects of income and income-related personal characteristics may pose significant problems for the transferability of results between countries.

In the final part a regression analysis is undertaken using the largest sample of air pollution mortality studies to date, from both developing and developed countries, in an attempt to further the understanding of the relationship between suspended particles and mortality. Applying empirical Bayes meta-analysis, it is estimated that mortality rates on average increase by 6 per cent per 100-μg/m³ increase in Particulate Matter (PM₁₀) concentrations, with greater effects in countries with high income inequality. We further find evidence that education and income have an influence on the effects of PM pollution.

This thesis shows that the traditional models and measurements of illness and of the association between pollution and health can lead to distorted estimates of illness in general, and of the adverse health effects of air pollution, in particular, by neglecting the interaction between income and health.
Table of Contents

Acknowledgements...................................................................................................................2
Abstract.................................................................................................................................3
Table of Contents....................................................................................................................4
List of Tables ............................................................................................................................7
List of Figures............................................................................................................................7
List of Abbreviations................................................................................................................8

Chapter 1: Overview and Summary..........................................................................................9
1.1 INTRODUCTION ............................................................................................................. 9
1.2 THE DISTRIBUTIONAL EFFECTS OF ILLNESS WHEN RECUPERATION IS ENDOGENOUS........ 10
1.3 WHY LEVEL AND DISTRIBUTION OF INCOME MATTER FOR THE HEALTH COSTS OF AIR
    POLLUTION..................................................................................................................... 12
1.4 CAN POPULATION CHARACTERISTICS ACCOUNT FOR THE VARIATION IN HEALTH IMPACTS OF
    AIR POLLUTION? A META-ANALYSIS OF PM10-MORTALITY STUDIES......................... 14
1.5 THESIS OUTLINE ........................................................................................................ 15

PART 1: THE DISTRIBUTIONAL EFFECTS OF ILLNESS WHEN RECUPERATION IS ENDOGENOUS

Chapter 2: Model Context and Relevance..............................................................................18
2.1 INTRODUCTION ............................................................................................................. 18
2.2 LITERATURE REVIEW .................................................................................................. 19
    2.2.1 Sickness Time ........................................................................................................... 19
    2.2.2 Welfare Calculations ............................................................................................... 22
    2.2.3 Interim Summary ..................................................................................................... 26
2.3 DESCRIPTION OF THE PROBLEM ................................................................................ 27
    2.3.1 Modelling Illness ................................................................................................... 27
    2.3.2 Efficient Time Opportunity Frontier ......................................................................... 28
    2.3.3 Why Work During Illness? ...................................................................................... 30
    2.3.4 Utility Maximisation ............................................................................................... 32
    2.3.5 Welfare and Income Effects of Illness ................................................................... 32
2.4 LIMITATIONS OF A ONE-PERIOD MODEL .................................................................. 34
2.5 SUMMARY ..................................................................................................................... 36

Chapter 3: Wage Rate and Recuperation Choice.....................................................................38
3.1 INTRODUCTION ............................................................................................................. 38
3.2 THE MODEL ASSUMPTIONS ....................................................................................... 38
3.3 EFFICIENT TIME OPPORTUNITY FRONTIER .............................................................. 39
    3.3.1 Regime 1: Part of Post-Recuperation Time Spent in Leisure ..................................... 40
    3.3.2 Regime 2: No Post-Recuperation Time Spent in Leisure ......................................... 42
    3.3.3 Interim Summary ................................................................................................... 47
3.4 UTILITY MAXIMISATION .............................................................................................. 48
    3.4.1 The Marginal Effect of Wage Upon Leisure .............................................................. 48
    3.4.2 The Existence of a Critical Wage Above Which People Choose to Fully Recuperate ... 51
3.5 SUMMARY ..................................................................................................................... 54
Chapter 4: Measuring the Adverse Effects from Illness .................................................. 57

4.1 INTRODUCTION ............................................................................................................. 57
4.2 EFFICIENT TIME LOSS (ETL) FROM ILLNESS AND ABSENCE FROM WORK (WA) .......... 57
   4.2.1 ETL and WA in the Full-Recuperation and the Exogenous-Recuperation Cases .......... 59
   4.2.2 Illness Case 1 ........................................................................................................... 61
   4.2.3 Illness Case 2 .......................................................................................................... 63
   4.2.4 Interim Summary .................................................................................................... 65
4.3 INCOME LOSS FROM ILLNESS ...................................................................................... 68
   4.3.1 Income Loss in the Full Recuperation and the Exogenous Recuperation Cases .......... 69
   4.3.2 Illness Case 1 ........................................................................................................... 73
   4.3.3 Illness Case 2 .......................................................................................................... 76
   4.3.4 Interim Summary .................................................................................................... 81
4.4 WELFARE LOSS FROM ILLNESS .................................................................................... 84
   4.4.1 Welfare Loss in the Full Recuperation and the Exogenous Recuperation Cases .......... 87
   4.4.2 Illness Case 1 ........................................................................................................... 89
   4.4.3 Illness Case 2 .......................................................................................................... 93
   4.4.4 Interim Summary .................................................................................................... 97
4.5 ADVERSE EFFECTS FROM ILLNESS – TWO EXAMPLES .................................................. 100
   4.5.1 Example 1 .............................................................................................................. 101
   4.5.2 Example 2 .............................................................................................................. 106
   4.5.3 Comparison of Income and Welfare Losses in the Two Examples ...................... 110
   4.5.4 Relevance ............................................................................................................. 113
4.6 SUMMARY ....................................................................................................................... 114
4.7 CONCLUSION .................................................................................................................. 117

PART 2: WHY LEVEL AND DISTRIBUTION OF INCOME MATTER FOR THE
HEALTH COSTS OF AIR POLLUTION

Chapter 5: Why Level and Distribution of Income Matter for the Health Costs of
Air Pollution ....................................................................................................................... 123

5.1 INTRODUCTION .............................................................................................................. 123
5.2 HEALTH IMPACT OF AIR POLLUTION – QUANTIFICATION, VALUATION, AND APPLICATION ..................................................... 124
5.3 THE MODEL ................................................................................................................... 129
   5.3.1 Health Production Tradition ................................................................................. 129
   5.3.2 A Clarification: Why the Role of Income May Vary with Health Event Specification .... 132
5.4 WHY THE DOSE-RESPONSE COEFFICIENT MAY BE AFFECTED BY INCOME ................................................................. 134
   5.4.1 Mitigating Measures .............................................................................................. 134
   5.4.2 Baseline Health .................................................................................................... 137
   5.4.3 Averting Activity and Exposure ........................................................................ 140
   5.4.4 Individual Behaviour and Measured Health Effects ........................................... 146
5.5 WHY WILLINGNESS TO PAY MAY VARY WITH INCOME ........................................... 148
5.6 FOUR INCOME-RELATED CHALLENGES AND SUGGESTED CORRECTIONS ......................... 153
   5.6.1 Omitted Variable Bias ......................................................................................... 155
   5.6.2 Transfer Challenge ............................................................................................. 160
   5.6.3 Covariance Challenge ......................................................................................... 165
   5.6.4 Information Bias ................................................................................................ 170
   5.6.5 Overall Measured Income Effect ....................................................................... 177
5.7 SUMMARY AND CONCLUSION ....................................................................................... 180
PART 3: CAN POPULATION CHARACTERISTICS ACCOUNT FOR THE VARIATION IN HEALTH IMPACTS OF AIR POLLUTION? A Meta-Analysis of PM$_{10}$-Mortality Studies

Chapter 6: Can Population Characteristics Account for the Variation in Health Impacts of Air Pollution? A Meta-Analysis of PM$_{10}$-Mortality Studies.........................186

6.1 INTRODUCTION.................................................................186
6.2 A SURVEY OF EXISTING META-ANALYSES OF THE MORTALITY FROM AIR POLLUTION......188
6.3 THE MODERATOR VARIABLES.........................................195
6.4 SAMPLE SELECTION, DATA, AND METHODOLOGY ..............202
  6.4.1 Sample ....................................................................202
  6.4.2 Data .................................................................206
  6.4.3 Methodology ............................................................213
6.5 RESULTS.................................................................216
  6.5.1 The MG Sample .........................................................216
  6.5.2 Full Sample ............................................................224
  6.5.3 Sensitivity Analysis ....................................................230
  6.5.4 Discussion ..............................................................239
6.6 CONCLUSION .............................................................250

REFERENCES.............................................................................253

Appendices..............................................................................277

APPENDIX 3.A: MATHEMATICAL NOTES..............................................277
APPENDIX 3.B: POTENTIALLY DIFFERENT PRODUCTIVITY IN WORK AND LEISURE ......................279
APPENDIX 3.C: CRITICAL WAGE DERIVED IN THE CASES WHERE $\sigma = 1$ AND $\sigma = 0.5$ ..........................282
APPENDIX 3.D: THE ASSOCIATION BETWEEN UNEARNED INCOME AND LEISURE .........................283
APPENDIX 4.A: DIFFERENTIATIONS ...............................................284
APPENDIX 5.A: DIFFERENTIATION OF DOSE-RESPONSE FUNCTIONS ..................................................289
APPENDIX 5.B: HYPOTHETICAL EXAMPLES ................................290
APPENDIX 6.B: MATHEMATICAL DERIVATION OF THE EQUATIONS IN SECTION 6.4 .........................293
APPENDIX 6.C: DO-FILE IN STATA TO MAXIMISE THE LIKELIHOOD FUNCTION ..................................295
APPENDIX 6.D: DATA ..........................................................297
APPENDIX 6.E: REGRESSION RESULTS FOR THE FULL SAMPLE USING OLS AND VWLS ...............303
APPENDIX 6.F: REGRESSION RESULTS FOR A RANDOM SAMPLE OF SAMPLE COMBINATIONS ............305
List of Tables

Chapter 4:
Table 4.1: Healthy scenario \((a = 1, \bar{w} = 2, \alpha = 0.5)\) .................................................................................................................. 100
Table 4.2: Illness case 1 \((\beta = 0.8, r = 0.3, s = 0.5)\) .................................................................................................................. 103
Table 4.3: Sickness scenario with exogenous recuperation \((r = 0.3)\) .................................................................................................................. 104
Table 4.4: Illness case 2 \((\beta = 0.5, r = 0.3, s = 0.8)\) .................................................................................................................. 108
Table 4.5: Comparisons of income and welfare losses in cases 1 and 2 .................................................................................................................. 112

Chapter 5:
Table 5.1: Summary of Estimation Biases related to Income .............................................................................. 180

Chapter 6:
Table 6.1: Summary of the expected signs of the moderator variables in a table .............................................. 201
Table 6.2: Summary of regression results using VWLS-methodology (MG-sample) .............................................. 220
Table 6.3: Weight given to the studies of the sample with VWLS (variance weighted least squares) ...................... 222
Table 6.4: Correlations between predictors considered in the restricted meta-analysis \((n=12)\). Values greater than 0.5 are in italics .................................................................................................................. 222
Table 6.5: Summary of regression results using EB-methodology (MG-sample) .............................................. 223
Table 6.6: Summary of regression results using EB-methodology on full sample \((70)\) ............................................. 227
Table 6.7: Correlations between predictors considered in the full meta-analysis \((n=70)\). Values greater than 0.5 are in italics .................................................................................................................. 229
Table 6.8: Summary of regression results using EB-methodology (sample excluding all-cause mortality studies \((61)\), sample excluding negative coefficients \((64)\), sample of TSP studies \((24)\), PMG studies \((33)\), and BS studies \((16)\)) .................................................................................................................. 236
Table 6.9: Summary of regression results using EB-methodology (sample of lag 0 coefficients \((28)\), lag 1 coefficients \((23)\), and lag 2 coefficients \((12)\)) .................................................................................................................. 237
Table 6.10: Summary of regression results using EB-methodology (sample consisting of one coefficient (averaged or actual) per country \((21)\), and taking time-period into account \((25)\)) .................................................................................................................. 238
Table 6.11: Data based on information given in studies .............................................................................................. 297
Table 6.12: Data gathered from SIMA and WHO .................................................................................................................. 300

List of Figures

Chapter 2:
Figure 2.1: Efficient time opportunity frontiers for a healthy period, and in the case of exogenous and endogenous sickness periods. .................................................................................................................. 29

Chapter 3:
Figure 3.1: Efficient time opportunity frontier for those who take some or all of post-recuperation time out in leisure .................................................................................................................. 42
Figure 3.2: Efficient time opportunity frontier in the case where exact full recuperation yields the maximum amount of efficient labour \((case\ 1)\) .................................................................................................................. 45
Figure 3.3: Efficient time opportunity frontier in the case where working during the entire recuperation time yields the maximum amount of efficient labour \((case\ 2)\) .................................................................................................................. 46
Figure 3.4: Efficient time opportunity frontier in the case where working part of the recuperation time yields the maximum amount of efficient labour \((case\ 3)\) .................................................................................................................. 47

Chapter 4:
Figure 4.1: Efficient time loss in the full recuperation and the exogenous recuperation cases. .................................................. 61
Figure 4.2: Efficient time loss in illness case 1. .................................................................................................................. 63
Figure 4.3: Efficient time loss in illness case 2. .................................................................................................................. 65
Figure 4.4: Work-absence at various wage rates for illness cases 1 and 2. .................................................................................................................. 67
Figure 4.5: Efficient time loss at various wage rates for illness cases 1 and 2, and for the EXOR-model. ...................... 67
Figure 4.6: Income losses (absolute and relative) at various wage rates for illness case 1 and the EXOR model. .......... 83
List of Abbreviations

ACWL  Absolute comparable welfare loss
AIL   Absolute income loss
BS    Black (British) smoke
C-D   Cobb-Douglas
CI    Confidence interval
DALE  Disability-adjusted life expectancy
DALY  Disability-adjusted life years
DIST  Distribution (interaction term of GNP per capita and the GINI-coefficient)
EB    Empirical Bayes
EDUC  Education
ENDOR Endogenous recuperation model
ETL   Efficient time loss
EXOR  Exogenous recuperation model
LEAB  Life-expectancy at birth
MG    Maddison and Gaarder (2001)
OVER65 Per cent of population over 65 years of age
PM    Particulate matter
POLL  Pollution
RCWL  Relative comparable welfare loss
RIL   Relative Income loss
TSP   Total suspended particles
WA    Work absence
WLD   Work loss days
WTL   Work time loss
WTP   Willingness to pay
Chapter 1: Overview and Summary

1.1 Introduction

Illness reduces market and non-market productivity, the total amount of time available for production, as well as individual well being. To avoid or minimise these unwelcome experiences, individuals can invest in their health in order to produce health, or at least restore part of it after an illness. These ideas were presented in an article by Selma Mushkin (1962), and formalised by Michael Grossman (1972). The literature on health capital has grown apace since Grossman’s seminal paper. The first part of this thesis aims to close a gap in the literature, with important distributional and illness-measurement implications. The allocation of sickness time to various activities (e.g. for the purposes of working or recuperation/leisure) may affect the duration of the current sickness experience, as well as the income and welfare losses incurred, once it is recognised that illness is a matter of degree and its impact on daily activities to a certain extent subjective.

The effects of behaviour on illness, be it through the decision to use medical care or through the decision to stay away from work and recuperate have not been fully integrated into the so-called ‘dose-response’ literature, as will be discussed in the second part of this thesis. Since the air pollution disaster in London in 1952, numerous empirical studies have observed associations between air pollution and various adverse health events. The information retrieved from such studies has been used to estimate and quantify the potential health benefits of controlling air pollution. To date, the possible interactions of income and illness through behavioural choices, and the implications for quantification and transferability of findings have been largely ignored.

The level of income may affect the amount of averting and mitigating activities undertaken, and the willingness to pay to avoid an adverse health effect. The validity of

---

3 Studies in which the effect of a change in air pollution exposure on the amounts of morbidity or mortality experienced is estimated.
4 The London smog disaster (December 1952) established that high levels of air-borne particles and sulphur dioxide produced large increases in daily death rates (HMSO (1954)).
5 See for example Dockery et al. (1993), Ostro et al. (1996), Anderson et al. (1996), Bacharova et al.
the methods used to quantify and value the health impact of air pollution is therefore questioned. Furthermore, the effects of income and income-related population characteristics may pose significant problems for the transferability of results between countries.

If income and income-related characteristics affect the association between pollution and mortality, we would expect to observe variability among epidemiological findings. In the final part of the thesis the largest sample of air pollution mortality studies to date, from both developing and developed countries, is analysed in an attempt to further the understanding of the relationship between particles and mortality. Previous meta-analyses addressing between-study variability have not included such a variety of population and income-related characteristics in their attempts to explain the differences in effect estimates.

This thesis shows that the traditional models and measurements of illness and of the association between pollution and health can lead to distorted estimates of illness in general, and of the adverse health effects of air pollution, in particular, by neglecting the interaction between income and health. The thesis consists of three interrelated but separately publishable parts. The next three sections, 1.2 through 1.4, present the aims, main research questions, methodologies, and findings of each of these in turn, and can be read as an executive summary. Section 1.5 describes the outline of the thesis.

1.2 The Distributional Effects of Illness when Recuperation is Endogenous

The main objective of the first part of the thesis is the analysis of endogenous recuperation and its effect on illness duration, and income and welfare distribution. In particular, the conditions under which an individual may decide to work rather than stay at home during the recuperation time from illness are examined. Two main and connected research questions are addressed: Does wage level affect labour supply behaviour under illness? And if so, how does this affect the path of the illness and thereby income, welfare, and illness measurements?

Standard utility maximisation theory is used to analyse an individual’s labour supply behaviour under constraints imposed by sickness and minimum consumption requirements. Departing from the previous literature, productivity in work and utility derived from leisure (productivity of leisure time) in this model are influenced by the labour supply decision: if the individual works during exogenously imposed full
recuperation time, productivity in the post-recuperation time suffers, and produces a non-linear efficient time budget constraint. We introduce a way of describing illness that is particularly useful for the case where the recuperation period is endogenous. Sickness is identified by three parameters: duration, severity in terms of productivity loss, and efficiency of recuperation.

For a rather general utility function, an empirically supported elasticity of substitution between leisure and consumption, and under the assumptions that no sick pay is received and that wage is paid according to productivity, we find that only higher-wage individuals will choose to recuperate fully, whereas others will work while being sick. In other words, poor people have to work while being sick because of a minimum consumption requirement. The main implication of this finding is that income and welfare losses from a given illness may be larger for the low-wage individuals than for those with higher wages. However, low-wage individuals choose to continue working during their illness only if they lose less income than if they had stayed at home to recuperate.

The usual way of modelling sickness as time lost (and assuming the time loss is equal across wage groups) would, in contrast, imply higher income and absolute welfare losses for higher wage groups. The traditional model may thus seriously misrepresent the income and welfare losses experienced due to illness.

A further implication is that commonly used statistical measurements of illness can lead to distorted illness-estimates by neglecting the interaction between income and behavioural choices. Four types of illness-measurements and the implication of their use in severe and less severe illness cases are presented.

Work-absence is shown to be a poor measurement of the adverse effects of illness because it does not capture the illness of those who work although being sick, and does not distinguish between imposed and chosen recuperation.

Efficient time loss is not well-suited as a measurement of illness effects as it does not take into account that the efficient time spent in leisure and labour may not be worth equally much to the individual.

Income loss as an adverse effect measurement captures the fact that in severe illness cases the loss of opportunity to earn additional income during the illness period is total, whereas for less severe illnesses it is only partial. However, the income loss measurement ignores other welfare aspects of being healthy (e.g. effect on leisure time).
Welfare loss is the correct measurement of the adverse health effect as it takes individual's preferences for leisure and consumption into account. By measuring the full income equivalent at a reference wage the losses are comparable over individuals.

In addition to adding to labour and health economics, the insights from the model development and analysis may have important policy implications, especially for developing countries. In particular, health and pollution policies may have redistributing effects. Both the welfare and the income losses due to a given illness are under certain conditions larger for the lowest wage groups than for those with higher wages, implying that the occurrence of certain common illnesses or illness-reducing policies (e.g. vaccination-projects, pollution-reduction initiatives) may have income and welfare redistributing effects.

1.3 Why Level and Distribution of Income Matter for the Health Costs of Air Pollution

The main goal of the second part of the thesis is the analysis of the possible interactions of income and illness through behavioural choices, and the implications for quantification, valuation, and transferability of findings from air pollution mortality studies. Two main and connected research questions are addressed: is the way in which the health impacts of air pollution have been quantified and valued in high-income countries correct? Even for correct estimates, what do the effects of income and other personal characteristics imply for transferability into developing country contexts?

The dose-response function is presented as a reduced form, which can be derived from a simple health production function. Hence, pollution has not only a direct effect on health through exposure, but in addition affects health indirectly through its influence on mitigating and averting activities. The latter activities reduce the exposure to air pollution, whereas the former reduce the health impact from pollution exposure. A new and potentially important modelling distinction is made between studying the health effect as the probability of falling ill (or of dying), and the health effect taking into account the severity of the illness (in terms of duration and/or impairment). The distinction is an important one to make as there is no reason why income should have the same effect on the probability of illness as it has on severity.

If we accept that income level may affect the amount of averting and mitigating activities undertaken by individuals, their willingness to pay to avoid an adverse health
effect, as well as their baseline health level, then there are four reasons why it is problematic to ignore the influences of income when estimating the health costs from air pollution.

(1) The first challenge is that of omitted variable bias. If the adverse health effects from air pollution depend on personal characteristics and individual income level, then failure to include these variables in the regression analysis will lead to the effect estimates being biased.

(2) However, even with unbiased estimates for one particular population sample, transferring an adverse health effect estimate onto another population with different characteristics will be incorrect. We suggest that one way of dealing with this transfer challenge is to introduce moderator variables, among them income, in order to tailor-make the dose-response coefficient to the country in question based on local characteristics. The coefficients used for these moderator variables are obtained by regressing estimated dose-response coefficients (from meta-analysis) onto these variables.

(3) Even with unbiased estimates for one particular population sample, however, the quantification of the health effect in that population will be incorrect as different income groups within the population have different willingness to pay (WTP) to avoid illness. The health costs are usually estimated by multiplying the estimated WTP and dose-response coefficient for the entire population under investigation, and any covariance between health effect and WTP is thus ignored. If the adverse health effects are larger for low-income groups, but the WTP lower, the covariance will be negative and poses a third challenge. The cost estimates of mortality from air pollution can be improved by combining the income-sensitive slope estimate mentioned above (point 2), and information on the income-sensitivity of the WTP.

(4) Finally, the traditional measurements of the health effect from air pollution may be particularly unsuited to capture the true effect on the poorer income groups due to differing behavioural responses to illness in different income groups. Improved health effect measures may be obtained using information on income loss, or through interviews/self-reported illness, and medical check-up. Similarly, the degree to which ambient pollution levels can be used as a measurement of exposure to air pollution may vary with income. The exposure measurements can be improved through greater spatial density of outdoor monitoring, indoor monitoring or information about what type of
heating and cooking fuels individuals are using, and with information concerning the amount of time an individual spends in each micro-environment.

There are a number of potentially important implications from this discussion. First, the failure to take the health effect of air pollution on different income groups properly into account may lead to wrong estimates of the health cost of air pollution, and therefore possibly to wrong decisions when carrying out benefit-cost analysis of pollution-decreasing policies.

Furthermore, the actual effect of pollution on health may vary with income, and therefore make transfers of estimates from high-income countries to low-income countries particularly undesirable. As long as the health measurements used are based on behavioural responses to illness rather than the illness itself, however, carrying out studies in situ may not necessarily lead to better estimates than those obtained by transfer. Furthermore, increasing the income of the poorest groups may lead to less actual adverse health effects, but a larger amount of measured health effects.

Even in the absence of pollution reductions through pollution control policies, the health effect of pollution in developing countries may be reduced through such measures as the introduction of paid sick leave, or through income redistributing measures.

1.4 Can Population Characteristics Account for the Variation in Health Impacts of Air Pollution? A Meta-Analysis of PM$_{10}$-Mortality Studies.

The aim of the final part of the thesis is to analyse the variation in estimated dose-response coefficients across studies. Applying mixed effect Empirical Bayes meta-analysis (Raudenbush and Bryk (1985)), we run regressions over the largest sample of air pollution mortality studies to date, from both developing and developed countries, in an attempt to further the understanding of the relationship between suspended particles and mortality. Previous meta-analyses addressing between-study variability have not included such a variety of population characteristics in their attempts to explain the differences in effect estimates.

The sample of estimated dose-response coefficients are regressed on relevant exogenous variables, also called moderator variables (income, income inequality, baseline health, education, per cent of population over 65 years of age, and health services), in order to investigate whether these can explain the variation among the dose-response
coefficients. We assume that the estimated effect size in each study is a function of known study characteristics, random errors (inter-study variability), and errors of estimation (intra-study variability). Assuming that the estimated variability from each study is approximately equivalent to its true value, we use maximum likelihood to determine the inter-study variability. A program was developed for this analysis in STATA (econometrics package) in order to maximise the above likelihood function. A detailed sensitivity-analysis is carried out.

We estimate that mortality rates on average increase by 6 per cent per 100-μg/m³ increase in PM_{10} concentrations, with greater effects in countries with high income inequality. We further find evidence that education and income have an influence on the effects of PM pollution.

In addition to income inequality possibly proxying for the quality and availability of health services, as well as other inequalities (e.g. in baseline health, education), its importance may indicate that the income level we would have wanted to measure is not average per capita income in the country, but rather that of the city in question and more particularly of the potentially vulnerable individuals in that city. Furthermore, stress due to inequalities, as well as decreasing returns to averting and mitigating activities could both explain a tendency for higher inequality to be associated with higher changes in mortality rate due to increases in air pollution.

Interestingly, we find that the results are rather sensitive both to the lag structure of daily air pollution measurements, and to the inclusion or exclusion of mortality from external causes in the mortality measure.

The results from the meta-analysis contribute to the understanding of what affects the amount of deaths related to air pollution, and helps highlight areas where further research is needed. Furthermore, the insights obtained can help guide policy decisions in the future. More particularly, the findings can be used to adapt dose-response coefficients to local conditions.

1.5 Thesis Outline

First chapter gives an executive summary and main conclusions of the thesis. The first part starts in chapter 2 and ends with chapter 4. Chapter 2 consists of a literature survey (section 2.2) and a description of the ideas and concepts that form the background of the model (section 2.3). The model is then developed in chapter 3, analysing the effect
of wage level on labour supply behaviour under illness. Specifically, the model assumptions are reviewed in section 3.2, whereas the model itself consists of two stages. First, the efficient time opportunity frontier is derived in section 3.3. Next, the optimal location on this frontier is derived for different wage groups by maximising utility in section 3.4. Section 3.5 summarises the main findings. Chapter 4 discusses different methods used to measure the adverse effect of illness, and compares the predictions of using these measurements for different wage groups. Efficient time loss and absence from work are covered in section 4.2, income losses in section 4.3, and welfare losses in section 4.4. Two simulation examples are made in section 4.5, and section 4.6 summarises the chapter. Section 4.7 concludes the entire first part of this thesis.

The second part is presented in chapter 5, and starts out by reviewing how the health impact of air pollution has been valued and quantified using epidemiological dose-response studies and willingness-to-pay (WTP) studies (section 5.2). The dose-response function is then analysed within the wider framework of health production functions in section 5.3, and modelling clarifications are made. Section 5.4 discusses the various reasons why the measured health effects from air pollution may vary between income groups, whereas section 5.5 looks at the effect income may have on the willingness-to-pay. Section 5.6 considers the problems posed by income and income distribution both to the quantification and cost estimation of the health impact of air pollution, and to the transfer of estimates from one population to another, and makes suggestions for improvements. The last section, 5.7, summarises the main insights and concludes with some methodological implications.

The third part is presented in chapter 6. Section 6.2 introduces the concept of meta-analysis, as well as the various uses, strengths, and weaknesses of this type of analysis, whereas section 6.3 reviews the moderator variables selected. Next, section 6.4 describes the criteria used in composing our sample of past studies, the data used to capture the moderator variables, as well as the model and estimation procedure. The main results are presented in section 6.5, together with a sensitivity analysis and a discussion of the findings. Main results and implications are reviewed in chapter 6.6.
PART 1: THE DISTRIBUTIONAL EFFECTS OF ILLNESS WHEN RECUPERATION IS ENDOGENOUS
Chapter 2: Model Context and Relevance

2.1 Introduction

Part 1 of the thesis, consisting of three main chapters, explores the issue of labour choice during illness and its effect on efficient time, income, and welfare distribution. Two main and connected questions are addressed: Does wage level affect labour supply behaviour under illness? And if so, how does this affect the path of the illness and thereby income, welfare, and illness measurements?

Standard utility maximisation theory is used to analyse an individual's labour supply behaviour under constraints imposed by illness and minimum consumption requirements. We assume that there is no sick pay and that the wage rate depends on productivity. Furthermore, departing from the previous literature, productivity in work and utility derived from leisure (productivity of leisure time) in this model are influenced by the labour supply decision: if the individual works during exogenously imposed recuperation time, productivity in the post-recuperation time suffers, and produces a non-linear efficient time budget constraint.

We find that for a rather general utility function and an empirically supported elasticity of substitution between leisure and consumption only high-wage individuals choose to recuperate fully from their illness before returning to work. Low-wage individuals choose to work part of the recuperation time because they lose less income than if they had stayed at home to recuperate, however, they thereby lose productivity in the post-recuperation time. Main implication of this finding is that under certain assumptions and parameter specifications income and welfare losses for the low-wage individuals may be larger than for those with higher wages.

The traditional way of modelling illness as time lost and commonly-used statistical measurements of illness effects, such as work loss days, are therefore shown to lead to distorted estimates of illness by neglecting the interaction between income and health.

The current chapter starts in section 2.2 with a brief literature survey. This is followed in section 2.3 by a description of the ideas and concepts that form the background of the model. A general style model is then developed in chapter 3, analysing the effect of wage level on labour supply behaviour under illness. Chapter 4 discusses
different methods used to measure the adverse effect of illness, such as efficient time, welfare, and income losses, and absence from work. Comparisons are further made between the different wage-groups, as well as between the endogenous and exogenous recuperation models.

We present a model where the labour decision is analysed for one period only. An extension to a multi-period model is left for future research, as are the issues of paid sick leave, health insurance, and medical investments.

2.2 Literature Review

In this section we review how sickness time and welfare are treated in the literature. Sub-section 2.2.1 presents a brief overview of the theoretical and empirical literature concerning sickness time, in particular the usage and endogeneity of sickness time. This overview is necessary in order to appreciate the main new feature of the model developed over the following chapters. Sub-section 2.2.2 discusses three different issues relevant for welfare comparisons: interpersonal comparisons of welfare, non-linearities in the budget-constraint, and welfare and the minimum consumption requirement. This section will address the main controversial issues relevant for the welfare loss comparisons carried out in chapter 3, and will expose where within the literature our treatment of the issues belongs.

2.2.1 Sickness Time

Illness reduces market and non-market productivity, the total amount of time available for production, as well as individual well being. These ideas were presented in an article by Selma Mushkin (1962), and together with Becker (1964) she introduced the concept of health capital, a component of the stock of human capital. The ideas were formalised by Michael Grossman (1972), who constructed a model of the demand for health capital. Grossman and others who have adopted his framework (e.g. Cropper (1977), Muurinen (1982), Wagstaff (1986)) consider sick time a function of health capital which in turn can be invested in and which depreciates due to various factors – this is known as the health-capital literature. Grossman argued that health capital differs from other forms of human capital in that a person's stock of knowledge affects productivity, whereas his or her stock of health determines the total amount of time available to be spent on producing commodities and earnings (Grossman (2000)). Whereas Michael Grossman and his followers used sickness time to investigate the demand for health and
health care, O'Donnell (1995) investigates the effects of sickness time on household time allocation and consumption decisions. This is at the outset similar to what is being done in this and the following two chapters. There are, however, important differences between the existing literature and the current analysis, and these will be reviewed below.

**Sickness time – time lost or time of choice?**

The theoretical literature has to date treated sickness time as time lost. Grossman (1972) assumes that the uses of healthy time in different activities are mutually exclusive and add up to total time available minus sick time. O'Donnell (1995) leaves Grossman's assumption unchallenged, while investigating how this given sickness affects household’s allocation of time and consumption. In a paper generalising the Grossman model, Muurinen (1982) points out that it is unrealistic to assume that individuals do not use medical care and invest in health during sickness, however, she does not explore the possibility and implications of an individual working during sickness-time.

Since neither the health-capital literature nor the model developed by O'Donnell allows sickness time to be used either for leisure or work, and because only healthy time is assumed to provide any utility, the severity of a sickness has therefore no role to play in their models.

In contrast to the theoretical tradition, the empirical literature has to a certain extent captured the idea that sickness time can be a time of choice. Thomas (1980) notes:

*Absence can be treated as a conscious decision on the part of the worker. With some forms of illness the individual may have no choice as to whether or not he attends work but with most absences he has discretion and at the margin it is legitimate to view absence as a labour supply decision.*

Similarly, Bartel and Taubman (1979) reveal in their article that sickness time has different potential usages;

*Some illnesses are only temporary and affect only current earnings. Other illnesses or their resulting impairments may persist for long periods. But a standard notion in economics is that inputs are substitutable in the production process. After a potentially persisting illness strikes, a person can undertake treatment to cure it, use other investments to overcome its effects, or move into an occupation in which it is less of a handicap.*

Paul Fenn (1981) attempts to show that the amount of time allowed for recuperation after illness is to some extent voluntary. He argues that sickness absence will

---

respond to changes in family structure, investment in training and education, and the
degree of income replacement, but will not respond to increased income replacement to
the same extent as the duration of unemployment will. In fact, in his empirical analysis he
finds that the elasticities of sickness duration to income replacement are 10% of those for
unemployment.

Finally, in a recent article ‘Sick but yet at work. An empirical study of sickness
presenteeism’ Aronsson et al. (1999) found that a third of the persons in their study
sample from Sweden had gone to work two or more times during the preceding year
despite being sick (subjectively measured).

Although these empirical studies reveal that choice over the usage of sickness
time is not a new concept, little theoretical attention has been given to date to the choice
an individual may face on how to spend his/her sickness time. Furthermore, none of the
above mentioned models and empirical investigations have analysed how that choice can
affect the duration of sickness time.

**Exogenous versus endogenous sickness time**

The health-capital literature has treated sick time as endogenous - a function of
health capital (inversely related) which in turn is affected by investments in health (health
care, medicines etc.) and depreciation of health due to factors such as use, age, disease,
and pollution. Individuals invest in health capital for two reasons. As a consumption
commodity, health may directly enter individuals’ preference functions, and as an
investment commodity it determines the total amount of time available for market and
non-market activities. The amount of health investment and thus sick time experienced by
an individual will depend on such factors as education and wage rate. However, this
literature does not explore how the usage of sickness time (e.g. for the purposes of
working, leisure-activities, or seeking medical care) may affect current sickness, both its
duration and severity. In other words, once the sickness sets in there is nothing that can
divert it from its predetermined path. Therefore, sickness time is endogenous until it
occurs – it then becomes exogenous.

In contrast to the health capital tradition, which is concerned with the allocation of
healthy time and commodities and how these affect the amount of sick time experienced,
O’Donnell’s model is analysing how a certain amount of sick time affects household’s
allocation of time and consumption. Sick time in O’Donnell’s model is exogenous, as it
does not allow the sickness period to be affected by the household’s allocation. The model developed in detail in chapters 3 and 4 will attempt to cover that important gap.

2.2.2 Welfare Calculations

Measuring welfare and welfare changes is central in the analysis of the costs to society from illness and in the analysis of optimal policy implications. The traditional approach relied on observed income or expenditure as a welfare indicator. A main reason for this has probably been the accessibility of data for these variables. However, as has been recognised in the more recent welfare literature these measures ignore the contribution of leisure to the standard of living. As noted by Preston and Walker (1992), this is an important shortcoming because leisure’s share in full income is likely to vary significantly across individuals and across time for significant groups in the labour market. Of particular interest in the context of our model is the fact that if the amount of leisure is correlated with income, welfare comparisons between income groups that do not take leisure into account will be systematically biased. The use of utility-based welfare measures has therefore lately become more common and accepted. The latter take into account the opportunity sets and relative prices facing individuals (or households), as well as variations in demographic characteristics.

Three important challenges relevant to the issue at hand emerge in connection with the utility-based welfare literature; firstly, the question is whether it is meaningful to base welfare evaluations on utility changes only (Sen (1984)), the second concerns non-linearities in the budget constraint (Preston and Walker (1992)), and the third looks at welfare measurement of poverty and how it differs from merely that of low income (Lewis and Ulph (1988), Ulph (1978)).

Interpersonal comparisons of welfare

Sen points out, firstly, that by looking only at utility changes different types of pleasure and pain are not distinguished. The sadist’s pleasure of torturing another human being may give as much utility as the disutility from pain of the tortured one. Second, he is very critical towards interpersonal comparisons of utility. In particular, he argues that people learn to adapt to their circumstances so that utility in fact is a relative concept and that we adjust our scale of content to the level of welfare we can expect to achieve:
The underdog learns to bear the burden so well that he or she overlooks the burden itself. Discontent is replaced by acceptance, hopeless rebellion by conformist quiet, and - most relevantly in the present context - suffering and anger by cheerful endurance. As people learn to adjust to the existing horrors by the sheer necessity of uneventful survival, the horrors look less terrible in the metric of utilities.\(^8\)

According to Ulph (2001),\(^9\) Sen's critique is precisely to the point when it comes to measuring happiness. He agrees with Sen in that people have a psychological propensity to adapt to situations, so when asked about degrees of happiness people do not express great extremes (apart from those whom we classify as mentally disturbed because they are excessively depressed or happy). However, he argues that one of the great contributions of Economics has been to sever the link between utility and happiness. There is therefore no reason why we are not capable of expressing extremely dire situations on a utility metric that reflects people's true position.

Although the issues raised by Sen are important to keep in mind when involved in welfare comparison analysis, neither the problem of adding up different types of pleasure and pain, nor that of interpersonal comparison arise in the model developed in the present and following two chapters. The individuals in this model are theoretical constructs assumed to have the same utility function, and indeed differ only with respect to the wage they are earning. Therefore, we assume that meaningful welfare comparisons can be made, and that these may be based on individual utilities.

**Non-linearities in the budget constraint**

Preston and Walker (1992) noted that most of the literature that has taken leisure into account when measuring welfare has simply treated it as an additional commodity, overlooking the fact that what distinguishes the analysis of labour supply from that of commodity demands is the nonlinearity associated with the budget constraint. Nonlinear budget constraints can arise under different circumstances such as nonlinear income taxes (Hausman (1985), Preston and Walker (1992)), non-linear household production functions (Pollack and Wachter (1975)), and non-linear wage-hours locus (Deardorff and Stafford (1976)).

In addition to the classic full income expenditure function, which measures the value of a consumer's consumption plus her leisure (which are equal to her endowment of

---

\(^9\) Personal communication.
consumption and time), Preston and Walker review a number of welfare measures that can be used in the context of non-linear budget constraints:

The unearned income expenditure function which gives the income at zero hours required to attain some utility level at a reference wage (and therefore differs from the full income measure by the value of the time endowment);

the consumption level (of the composite commodity) required to attain some utility level at a given reference wage;

the consumption level required to attain some utility level at a reference level of hours of work;

the wage required to attain some utility level at a reference level of unearned income;

and, finally, the consumption level required to attain some utility level at a reference level of unearned income.\footnote{Preston and Walker (1992), p.2}

The present analysis will add another reason for non-linearities in the budget constraint that does not involve constraints imposed from outside such as taxation, limited working hours, or compatibility between inputs into production. Here, it is constraints imposed from inside a person due to minimum consumption requirements and decrease in productivity during illness that create the non-linearity. The measure of welfare used in this analysis is taken from the welfare measurement methods for non-linear budget constraints described in Preston and Walker (1992). The analysis draws on King (1983) by deriving measures of welfare from equivalent incomes (full income required to attain a certain level of utility) at a reference wage.

\textit{Welfare and the minimum consumption requirement}

There are several ways to model the minimum consumption constraint; it could be considered a relative or an absolute constraint, and it could be modelled as an endogenous or an exogenous constraint. The option of a relative constraint may be interesting both from a developing and developed country point of view, whereas the option of an absolute constraint is especially relevant when applying the model to developing countries.

One can think of the minimum requirement as \textit{endogenous} and varying with past and present choices made, which may vary with e.g. past and present income. Some past choices may for all practical purposes be viewed as irreversible, such as a child’s diet influencing how tall he will grow up to be and thereby his minimum nutritional requirement at a later age. Other past choices are at least slow to adapt to a change in the
circumstances, e.g. buying a house and having a mortgage to pay. Furthermore, the present choices may also affect the minimum consumption requirement; e.g. more hours of work may require a higher minimum intake of calories. An *exogenously* given constraint, not influenced by any choices, past or present, is perhaps a less realistic assumption but is a simplification often used both empirically and in theoretical models.

The minimum requirement can be considered a *relative* requirement, depending on for example the social class or income group one belongs to, or wants to belong to. Better-paid jobs may often require the employee to wear a certain type of clothing and perhaps entertain people in their homes. Similarly, in order to belong to a certain social group it may be necessary to dine at the 'right' restaurants, drive the 'right' sort of car etc. Losing ones status within such a group may be something that an individual would like to avoid at all costs - even at the cost of health.

An alternative way of thinking about the minimum consumption requirement is in terms of an absolute minimum requirement, often called a poverty line (Lipton (1997), Creedy (1997), and Lewis and Ulph (1988)). The idea behind such a poverty line is that individuals are substantially worse off as a result of being in poverty compared with being just above the poverty line. Hence, one could make each individual’s utility function depend on a threshold consumption or poverty level and falling below the line implies a strong and discrete drop in utility. Empirically (e.g. World Bank (1993), (1994), (1998/99), World Development Report), this poverty line is taken to be identical for all individuals. A criticism which could be made against such a procedure is that individuals vary in their requirements, and one could hence expect the utility drop to set in at different levels of consumption for different people, i.e. the poverty line is to a certain extent endogenous.\(^\text{11}\)

In the model presented in chapters 2 and 3 the minimum consumption requirement will be assumed as *exogenous and the same for all individuals* (although the working time needed to fulfil this requirement will be endogenous and, in particular, will vary with wage). For the sake of simplicity, the *net* minimum consumption requirement (i.e. the

\[^{11}\text{The poverty line may, according to Lipton (1997), be set either ‘(i) by the food energy method (FEM), i.e. where dietary energy intake per adult equivalent has an expected value just enough to meet requirements; or (ii) by the purchasing-power parity or PPP method...’ (Lipton (1997), p. 1003). However, Lipton points out that the energy requirement method abstracts from variations in needs, variations which partly may be caused by poverty itself. Adult height and weight will in general vary from one individual to the next, and in particular will be lower among poor groups. When it comes to the PPP method, measuring purchasing power over the consumption bundle in one country is problematic if the consumption components and the relative prices of poor people’s consumption bundles are different.}\]
minimum consumption requirement net of non-wage income) will also be assumed the same for all individuals in the numerical examples. Given the first assumption of equal minimum consumption requirement for all, this second assumption implies the same level of unearned income for all individuals. Endogenising the minimum consumption constraint could be an interesting extension of the model, as would allowing for differences in unearned income.

The interpretation of the minimum consumption requirement as an absolute minimum existence requirement will be used in modelling labour supply behaviour under sickness. The individual will only have a choice over the time that is over and above the minimum labour time necessary to keep alive, i.e. the minimum consumption requirement on wage earnings has to be fulfilled. This is reflected in the utility function by subtracting the amount of hours necessary to fulfil this requirement from the amount of working hours where the individual possesses a choice. This does not imply that being alive does not bring utility, and thereby some amount of welfare, but rather that it is a precondition for choice. Furthermore, it can be argued that the welfare achieved by just being alive is the same for all individuals, or at least should be treated equally, in a welfare comparison between living individuals.

### 2.2.3 Interim Summary

The model developed in chapter 3 will expand upon the concept of endogeneity of illness introduced by Mushkin (1962) and Grossman (1972) by making current sickness dependent on current labour decisions, rather than on previous health investments. It will furthermore, in line with O'Donnell (1995), investigate the consequence of illness on time allocation, but will introduce a more realistic two-way causality by allowing time-allocation to also affect illness. The welfare estimates will be based on those developed by King (1983) and Preston and Walker (1992) that are concerned with non-linear budget constraints. However, the non-linearity of the budget constraint will in contrast to most non-linearities discussed so far in the literature be due to constraints imposed from within rather than from outside the individual. Finally, the minimum consumption requirement will be treated as an exogenous poverty line, in line with Creedy (1997), and Lewis and Ulph (1988).
2.3 Description of the Problem

In this section the essential features of the model will be discussed in a non-mathematical way in order to expose the main ideas in an accessible manner. First, subsection 2.3.1 explains how sickness will be modelled. The shape of the efficient time opportunity frontier is then described in sub-section 2.3.2. This is followed by a discussion in sub-section 2.3.3 of the reasons for working during illness. Based on the discussions of the previous two sub-sections, sub-section 2.3.4 looks at where on this line individuals will choose to locate their choice of efficient leisure and efficient labour. Finally, the potential welfare and income effects of illness are discussed in sub-section 2.3.5, and sub-section 2.3.6 summarises the main insights.

2.3.1 Modelling Illness

An individual may be either sick or healthy, and a period of time for this individual may correspondingly be divided into sick time and healthy time. Healthy time can be thought of as the time in which a person performs all activities at his or her highest productivity level, and derives the most utility out of leisure time. Sickness time, on the other hand, will be identified by lower productivity and lower utility derived from leisure than in the healthy situation.

Two terms are closely connected to sickness time and need to be defined: full recuperation time and actual recuperation time. Full recuperation time is the time it takes for the individual to regain full productivity given that the entire sickness time is spent not working. Actual recuperation time is the amount of time that the individual chooses to stay home and recuperate – this will be smaller or equal to full recuperation time.

Sickness will be identified by three parameters: the duration of the sickness, the severity of it, and, finally, the efficiency of the recuperation time.

1. Duration of sickness: The choice of labour supply will determine the duration of the sickness period. If the person recuperates fully, then the duration of the sickness will be the same as the full recuperation time. If, however, the individual should choose to work part of the full recuperation time (i.e. the actual recuperation time is shorter than full recuperation time), then this will prolong the duration of sickness time (low-productivity time).

2. Severity of sickness: Severity will be modelled in terms of productivity loss in both work and leisure.
3. The efficiency of the recuperation time: The idea behind recuperation time is that some uses of time will be more beneficial for recovery from an illness than others. In particular, we are in this model going to assume that staying at home (i.e. away from work) will speed up the process of recovery. The efficiency will therefore depend on the time actually spent recuperating, as well as on the severity of the illness.

Other things being equal one would typically expect that the longer the duration of the full recuperation period the more likely that the individual has to work during illness in order to fulfil his or her minimum consumption requirement. Also, other things held constant and assuming a wage based on productivity, the more serious the illness the less reason for working during the sickness period as less will come out of it in terms of additional consumption. Finally, the higher the efficiency of the recuperation time (cf. point 3 above), the less likely the individual is to work during this period.

The relationship between severity, duration, and productivity of recuperation will typically vary for different types of illnesses, and for different types of work.

In our model, a purely exogenous probability of falling ill is considered. The health shock is not affected by individual actions but, in contrast, the development of the illness is. The reason for modelling it this way is two-fold: first of all it may be an acceptable approximation for certain types of health shocks (e.g. an influenza-epidemic, or for the health effects of an increase in air pollution). More importantly, however, the model is intended to complement the health-as-capital models that show that the level of health capital may vary with income; the current model shows that even for the same amount of potential illness experienced (and implicitly the same amount of health capital) low income individuals may experience higher welfare and income losses than higher income individuals due to the decisions they are forced to make.

2.3.2 Efficient Time Opportunity Frontier

With the insights from the previous paragraph we can now have a brief look at the efficient time opportunity frontier, and how it differs when an individual is healthy the entire period and when he or she is ill part of the period.

Let us consider efficient time, which is the full productivity time-equivalent (i.e. time available multiplied by the level of productivity), and assume that the productivity-level under sickness is less than under healthy conditions. In a healthy period, the choice of time-use does not affect the productivity level, and time can therefore be allocated to
either work or leisure with total efficient time staying constant. The line HH' in figure 2.1 shows this individual’s opportunity set.

When an individual is sick part of the period, on the other hand, productivity will be lower during the sickness period and this will decrease the total efficient time available. Hence, the efficient time opportunity frontier during a period with sickness will lie to the left of HH'. If sickness time was considered as time lost, i.e. a time used merely to recuperate with recuperation not providing any form of utility as in Grossman and O'Donnell’s models (discussed in section 2.2), then choice would only be available over the use of the remaining healthy time during the period. Since healthy time in our model is not used for health investments, in contrast to the Grossman tradition, time use would not matter for total efficient time and the budget line would be parallel to HH' (denoted by SxS'x in figure 2.1).

Figure 2.1: Efficient time opportunity frontiers for a healthy period, and in the case of exogenous and endogenous sickness periods.

However, in the model developed here sickness time is not assumed entirely lost. In particular, it is assumed that whether sickness time is spent recuperating at home or continuing working has an effect on the duration and severity of the illness, and that sickness time spent recuperating provides a certain amount of utility in terms of leisure. If the full recuperation time is spent recuperating, then we know that the post-recuperation time will be fully productive. Hence the loss in efficient time occurs only during the recuperation period in which the individual derives less than full leisure-utility, we will
call this the *direct* effect. For all choices of leisure and labour where the individual chooses to fully recover, the efficient time opportunity set will therefore be an inwards-shifted version of the healthy one. This is illustrated by line segment S\_N\_C in figure 2.1.\textsuperscript{12}

If part or all of recuperation time is spent working there will be an *indirect*, as well as a direct effect on the amount of total efficient time. The direct effect is again the reduced productivity in any activity during the full recuperation time. The indirect, or endogenous, effect comes from the reduction in post-recuperation period productivity due to not having recuperated fully. This effect will depend on the extent of actual recuperation, as well as on the severity of the sickness, and the efficiency of recuperation time in producing post-recuperation healthy time. Hence, the amount of total efficient time will depend on how much of the recuperation time the individual chooses to work.

The more an individual works during recuperation the less overall efficient time will he or she have. Because working during sickness has the additional negative effect of reducing post-recuperation productivity, an individual will always start by taking the recuperation time out in leisure (assuming a flexible work situation allowing the individual to take the full recuperation period out in leisure and to work long hours when healthy). Any person who is planning to take an amount of leisure that is at least as long as the full recuperation time will take the recuperation period out in leisure.

With these insights we can now derive the efficient time opportunity frontier for a period that involves some sickness. For efficient leisure time at least as large as efficient recuperation time, the opportunity frontier under sickness is just a parallel inward shift of the healthy opportunity frontier. For efficient leisure time less than efficient recuperation time, the more of the recuperation time which is spent working the less total efficient time i.e. the time opportunity frontier is convex to the origin. This is illustrated by line C'CS\_N in figure 2.1.

### 2.3.3 Why Work During Illness?

Investments in health, whether in terms of taking time to recuperate or buying medicines, may affect the development of an existing health effect. Although a sickness would under certain conditions be of a certain severity and length, these may to an extent be changed by own decisions. A person may for example decide to work although being ill, and this may increase the time-period in which productivity and welfare are lower.

\textsuperscript{12} The line segment is not shifted as much inwards as was the case under the assumption of sickness-time
The reasons for working instead of recuperating can be several:

1. Without sick pay and with insufficient savings or credit available, the lost earnings may weigh higher than the prolonged recuperation time and lost leisure time;
2. With an uncertain job-future (e.g. due to temporary contract) an individual may decide to work during illness in order to increase the probability of future contracts;
3. The relationship between recuperation and illness is uncertain and may vary with the type of illness, as well as type of person and work (e.g. blue- versus white-collar work);
4. The type of work the individual is doing may leave him or her with no choice (objectively or subjectively felt) but to work, e.g. if it is a position of high responsibility, or due to irreplaceability.

One would expect the choice over labour supply under sickness in different income groups to differ. First, in the absence of sick pay arrangements, poor people may have to work although they are sick in order to fulfil their minimum consumption requirements. This argument will be particularly relevant in developing countries where the social welfare system often is insufficient and employees are not offered paid sick days.

There are three additional reasons why the labour-supply choice under sickness may differ between income groups. First, high responsibility jobs (referred to under point 3 above) may arguably tend to be held by non-poor individuals. On the other hand, Aronsson et al. (1999) found empirically an increased risk of being at work when sick for occupational groups where there are difficulties in finding replacement or a stand in, and these were often occupations with relatively low pay (e.g. teachers, nurses). Second, there may be a tendency for different income groups to hold different types of work (e.g. blue-collar versus white-collar), although the trend is not clear. It is arguable that certain types of sickness have a larger effect on productivity in more physical types of work, although others may have just as large an impact on mental abilities. In addition, the type of work performed during the sickness period may determine the degree to which the recuperation time is extended. Finally, it is possible that different income groups tend to hold different types of contracts (e.g. permanent versus fixed term). A recent empirical study (Benavides et al. (2000)) found that fixed term contracts and temporary work were negatively associated with absenteeism.

being time lost, because some leisure-utility is now derived from recuperating.
In the model developed next we will disregard these three latter possible differences between income groups because they are less clear-cut, because they require modelling of specific industries, jobs, contracts, and illnesses, and most importantly because the main focus of the model are developing countries where the first argument arguably is of most relevance.

2.3.4 Utility Maximisation

Knowing the efficient time budget constraint, we must find out where on this line a person chooses to locate his or her choice of leisure and labour. If we assume that all individuals have consumption and leisure objectives, and that the former can be interpreted as an income target, then if all individuals have the same preferences they would all choose the same bundle of labour and leisure when facing the same efficient time budget constraint. However, we assume that people have a minimum consumption requirement on wages and this will restrict the amount of efficient time over which the individuals have a choice. The minimum efficient working time requirement implied by a given minimum consumption requirement will vary with the level of the individual’s wages. For higher wages less working hours are required to fulfil the requirement ceteris paribus. Hence, poorer (low-wage) workers will have a more restricted choice than richer workers. If the minimum required efficient working hours of the poor are at a sufficiently high level, they may be forced to choose a bundle on the concave part of the efficient time opportunity boundary. In other words, they may have to work although they are sick in order to survive. This will mean that these people have less total efficient time, as well as less choice time (i.e. total efficient time less the time used to fulfil the minimum consumption requirement) to be used in the production of utility.

2.3.5 Welfare and Income Effects of Illness

The next question that needs to be addressed is what the adverse health effects imply for welfare and income.

If a person derives utility from both consumption and leisure and they are both normal goods, and if sickness affects both the productivity at work (and thus the amount of consumption obtainable) and the amount of utility derived from leisure time, or any one of these, then sickness will obviously affect the welfare of the individual. Differential adverse health effects will therefore have differential welfare effects. However, for a more detailed discussion further functional assumptions and parameter specifications are
required and we will return to a discussion of this issue later (see chapter 4).

How would we expect sickness to affect labour supply? An assumption fundamental for this discussion is that individuals are being paid according to their productivity, i.e. the wage is paid per efficiency hour and not per nominal hour. If we further assume that working during sickness prolongs the time period in which productivity is low, then three cases may be distinguished: i) no effect from sickness on labour supply, ii) a nominal time effect, and iii) an efficient time effect.

For some people sickness may not have any effect on labour supply because all the sickness time is taken out of leisure time. Since work and leisure tend to come in a certain regular order, whereas sickness is something for which a specific time of day cannot be allocated, this scenario requires a rather flexible work situation. Also, it requires that the person spend a certain amount of time in leisure usually, so that the sickness time does not need to "steal" some of the time allocated to working. One would probably expect a flexible work situation coupled with the choice of spending a certain amount of time in leisure to be characteristic of individuals who are in higher-wage employment.

If sickness does have an effect on labour supply, it may have so in two ways. First, it may decrease the amount of hours worked by the individual. Second, it may decrease the productivity of the individual and thus decrease the hours worked when counted in efficiency-units. In the first case one would expect sickness time to be longer than the usual chosen leisure time. We could also expect to see this situation for the type of jobs where the hours are less flexible. Most importantly, the person must be able to "afford" this loss of labour-time. In the last case, where labour in efficiency-units declines whereas the labour time in normal time units may stay the same or increase, the situation is rather different. Here we are dealing with individuals who do not have the choice to take sickness time out of leisure time because their amount of leisure time does not suffice to recuperate. However, they can not take it out of labour time either, because their earnings will then not suffice to satisfy their minimum consumption requirement. This gives them no alternative but to work during their sickness time, but with a lower productivity.

From the previous discussion it is clear that labour supply behaviour, and therefore income, may vary between income groups. Differential behaviour during

\[13\] As long as we assume consumption to be a normal good, efficient labour supply would not increase with
sickness induced by differential wages may therefore have an effect on income
distribution. For a more detailed discussion further functional assumptions and parameter
specifications are once again required and we will return to this issue later (see chapter 4).

2.4 Limitations of a one-period model.

The model used here describes a time allocation between different activities
within one period. Two aspects of the model are open to criticism: a) the (implicit)
assumption of a utility discount rate of 0 for within-period time allocation; and b) the
absence of inter-temporal linkages, i.e. the complete disregard the economic actor is
assumed to have for the world outside the one period. We will here provide some
arguments in support of the relatively simple structure chosen, and provide indication on
where extensions to the model might lead.

To answer the first issue, it should be noted that the model developed here
covers a relatively short time period in which the individual experiences a certain amount
of illness. Given the limited-period time-frame, we can assume that the effect of
dISCOUNTING would be small and therefore can be disregarded. Moreover, the fact that the
low-wage individuals in our model tend not to recuperate fully and therefore are not in
full health in the post-recuperation period does not reflect a difference between wage-
groups in the willingness to sacrifice current utility for future utility. What drives the
model is the minimum net consumption requirement on all individuals, and investing in
health (recuperating) is not compatible with living longer for the low-wage individuals
because their bare survival necessities are not fulfilled.

This said, if within-period utility discounting were introduced, this can be done
either based on the assumption of a constant or an endogenous discount rate. It has been
suggested that lower income groups, because of circumstance and environment, have
lower ability to defer gratification (Maital and Maital, 1978). In addition, one could
envision an extension of our model where e.g. the wage rate is endogenous and
determined by prior investments in human capital, and therefore low-wage individuals
would be those with higher utility discount rates. Applied to our model, which could de
seen as describing a world with two sub-periods, a lower wage rate could therefore be
associated with a higher utility discount rate. The less the individual cares about
tomorrow, the more he will be inclined to work today (during recuperation sub-period), as he will care less about recuperating in order to be healthy tomorrow, and more about obtaining some consumption also in the first sub-period. Hence, inclusion of time discounting could possibly reinforce the result that illness leads to increasing inequality.

Regarding endogenous determination of the discount rate, this would be driven by the linkage between current labour supply and second period health. In a paper by Becker and Mulligan (1997) the endogenous determination of time preferences is analysed. The authors note that individuals may alter their time preference in part by time and effort spent in forming mental pictures of future pleasures. One could think of extensions to our model where time spent recuperating could simultaneously be spent on anticipation, which again would imply that higher-wage individuals have a lower discount rate.

A proper analysis of these issues however is beyond the scope of this paper, and would have to take into account the possible effect on the discount rate of the linkage between current labour supply and second period health.

The more compelling issue concerns the absence of inter-temporal linkages in our model. On the one hand, one can think about how the decision on the extent of recuperation today affects the health stock tomorrow and so, presumably, (i) the probability of falling ill tomorrow; and (ii) the ease of recuperating tomorrow conditional on falling ill.

In particular, a realistic assumption would be that sickness is positively serially correlated, such that an increase in first period sickness due to the individual labour-choice raises the expectation of second period illness. This would lead to interesting contradictory effects of current sickness on current hours of work in a two-period framework; an increase in current period sickness has a reducing effect on current efficient labour supply whereas an expectation of future sickness has an increasing effect.

On the other hand, a proper inter-temporal theory would also have to look at savings and labour supply decisions. By not returning to work today the individual will accumulate less savings or, alternatively, more debt, which could intensify the pressure to

---

increase their chance of future survival, which would imply lower discount rates.

16 They cite as examples of such time and effort activities like acquiring information through schooling, access to print media, and time spent with older persons, particularly parents.

17 The literature that has dealt with the inter-temporal demand for health capital was briefly discussed in the literature review in section 2.2.

18 According to O'Donnell (1995) empirical evidence reveals a negative effect of sickness on labour supply, suggesting the direct effect of current illness outweighs the one through the expectation of future sickness.
return to work early tomorrow if one did fall ill. Furthermore, the decision not to return to work today could affect the probability of being laid off or getting employment tomorrow. Equally, if there is learning, the decision to stay home and recuperate could lower productivity tomorrow. Hence there are many complex inter-temporal linkages between decisions today and stocks of health, financial assets, and human-capital assets. It is therefore not at all clear that introducing inter-temporal considerations would lead to greater pressures to recuperate.

Any proper analysis of these issues has to be infinite horizon. A finite horizon model – say a 2-period model - always has a last period. This last period is essentially a one period model just as the one presented in this paper and it is difficult to know just how sensitive the conclusions in the early period are to the fact that there is an ultimate final period. A proper infinite horizon model that also models the within-period recuperation decision, however, would be intractable analytically.

2.5 Summary

Little attention has so far been given in the literature to the choice an individual may face on how to spend his or her sickness time and the effect of that choice upon the duration of sickness time has not been analysed. We have presented a way of describing illness that is particularly useful for the case where the recuperation period is endogenous. Sickness is identified by three parameters: duration, severity in terms of productivity loss, and efficiency of recuperation.

We have argued that although working during illness may affect the path of the illness and thereby the efficient time available to an individual, poor people may have to work although they are sick in order to fulfil their minimum consumption requirements. Furthermore, we have seen that sickness may affect both income and welfare, and have hence argued that differences in labour supply behaviour under sickness due to variations in wages may cause differences in the duration of a certain sickness and therefore differential income and welfare effects from sickness.

In chapters 3 and 4 this preceding argument will be presented in a formalised manner. The model presented in chapter 3 will consist of two main steps. First, the efficient time budget constraint is derived by calculating the maximum amount of efficient labour obtainable for a given amount of efficient leisure. Second, the bundle chosen on this time-budget constraint is derived. This is done by maximising utility
subject to the budget constraint and the minimum consumption requirements. We find that under certain assumptions wages and leisure are positively related.

The efficient time, income, and welfare losses from a certain illness, as well as the absence from work for different wage groups will then be analysed in chapter 4. A comparison will be made between the findings of the endogenous and the exogenous (traditional) recuperation models.
Chapter 3: Wage Rate and Recuperation Choice

3.1 Introduction

The main model, which also forms the basis for the analysis of income and welfare losses from illness in chapter 4, is developed in this chapter. In particular, we will address the question of whether wage level affects labour supply behaviour under illness. Using standard utility maximisation theory to analyse an individual’s labour supply behaviour under constraints imposed by sickness and minimum consumption requirements, we will show that only higher-wage individuals will choose to fully recuperate, i.e. not work during their illness. Chapter 4 addresses the question of how this result affects the path of the illness and thereby income and welfare.

First the main assumptions of the model are set out in section 3.2. Section 3.3 proceeds by deriving the efficient time opportunity frontier, which is essentially done by maximising efficient labour for a given amount of efficient leisure. Having derived the efficient time budget constraint, sub-section 3.4 addresses the issue as to where on this line utility maximising individuals with different wages will choose to lie. The chapter conclusions are given in section 3.5.

3.2 The Model Assumptions

The model uses standard utility maximisation theory to analyse an individual’s labour supply behaviour under constraints imposed by sickness and minimum consumption requirements. We assume that there is no sick pay and that the wage rate depends on productivity. Departing from the previous literature, productivity in work and utility derived from leisure (productivity of leisure time) in this model are influenced by the labour supply decision: if the individual works during exogenously imposed recuperation time, productivity in the post-recuperation time suffers.

We are assuming that there is 1 unit of time, and that any illness will occur at the beginning of this one-unit time-period. If a person falls sick it takes a fraction of the time \( r \) \((0 < r < 1)\) to recuperate, given that the person spends the entire recuperation time at home. We further assume a positive and identical net minimum consumption requirement on income for all individuals, i.e. the minimum consumption requirement is larger than non-wage income by a constant amount (either unearned income and the minimum
consumption requirement are the same for all individuals, or if one is higher for an individual then the other one is higher by the same amount). The individual has to decide how much of $r$ to stay at home and how much to work, knowing that working during the recuperation time will prolong the period in which the person is sick. Let $x$, $0 \leq x \leq 1$, be the fraction of $r$ that is actually spent recuperating (here: staying at home). Assume that the illness reduces the value of time in terms of productivity in work and leisure. Work-time in the recuperation phase (i.e. $r(1 - x)$) has productivity level $s$, and each hour of $r$ that is spent recuperating is regarded as equivalent to $s$, $0 < s < 1$, hours of leisure.\footnote{The assumption of equal productivity in work and leisure is not very important for the results obtained, and simplifies the exposition of the model. As long as the ratio of the productivity of work to the utility derived from leisure stays the same in the recuperation and the post-recuperation phase, the same results will be obtained. To see this refer to appendix 3 B.}

The rest of the time, $1 - r$, can also be spent in leisure or work. The effectiveness of post-illness time is affected by the choice of $x$, i.e. labour supply behaviour during recuperation time. Let $z$, $0 \leq z \leq 1$, be the fraction of the post-recuperation time spent in leisure (i.e. $z(1 - r)$), and let each hour of leisure be equivalent to $\sigma(x)$, $0 < \sigma(x) \leq 1$, hours of efficient leisure and each hour of work also have productivity $\sigma(x)$. Assume that if the entire recuperation phase is spent recuperating then the post-recuperation phase gives full productivity in work or leisure, i.e. $\sigma(l) = 1$, whereas if the full recuperation time is spent working the productivity in the post-recuperation time will be the same as during recuperation, i.e. $\sigma(0) = s$. This is another way of stating that if an individual does not take any time to recuperate, the sickness will in effect last the entire period.

For any given choice of $x$ and $z$ the consumer will get efficient leisure:

$$L = sxr + \sigma(x)z(l-r) \quad (3.1)$$

This takes the minimum value $L = 0$ when $x = z = 0$, and a maximum value $L = (l-r)sr$, when full leisure is taken and $\sigma(l) = 1$. Corollary to this choice of $x$ and $z$ is the amount of efficient work:

$$H = s(l-x)r + \sigma(x)(l-z)(1-r) \quad (3.2)$$

3.3 Efficient Time Opportunity Frontier

Assume that the consumer only cares about leisure and consumption. This implies that for any given amount of leisure, the consumer would want to get the maximum consumption possible and thus the maximum amount of efficient work $-\partial(L)$. The amount of efficient time, and the allocation to efficient work and efficient leisure, will be
a function of the choice of actual recuperation during the full recuperation time, \( x \), and of
the choice of leisure in the post-recuperation time, \( z \).

Define the maximisation problem in the following way:

\[
\Phi (L) \equiv \text{MAX } s \left( 1 - x \right) r + \sigma (x) (1 - z) (1 - r)
\] (3.3)

s. t. \( s x r + \sigma (x) z (1 - r) \geq L, 0 \leq x \leq 1, 0 \leq z \leq 1 \)

The solution to this problem determines \( x \) and \( z \) given the value of \( L \) and thus the amount
of recuperation a consumer will choose - conditional on \( L \). We form the following
Lagrangean:

\[
Z = s (1 - x) r + \sigma (x) (1 - z) (1 - r) + \lambda (s x r + \sigma (x) z (1 - r) - L) + \nu (1 - x) + \lambda (1 - z)
\]

Hence, the maximisation problem gives the following first order conditions (FOCs):

\[
\frac{\partial \Phi}{\partial x} = (\lambda - 1) \sigma (x) (1 - r) z + \lambda (1 - r) z \leq 0 \quad \text{and} \quad \lambda \sigma (x) (1 - r) z \geq 0
\]

\[
\frac{\partial \Phi}{\partial z} = (\lambda - 1) \sigma (x) (1 - r) z \leq 0 \quad \text{and} \quad \lambda \sigma (x) (1 - r) z \geq 0
\]

\[
\frac{\partial \Phi}{\partial \lambda} = s x r + \sigma (x) (1 - r) z \geq L
\]

We start by analysing the case where \( z > 0 \), i.e. where part of the post-recuperation period
is spent in leisure, which from here onwards will be called 'regime 1'. The second type of
regime, 'regime 2', occurs when no leisure is taken during post-recuperation (i.e. \( z = 0 \)).

3.3.1 Regime 1: Part of Post-Recuperation Time Spent in Leisure

Within regime 1 two sub-cases can be distinguished; in regime 1a the entire post-
recuperation period is spent in leisure, whereas in regime 1b only part of this period is
spent in leisure.

**Regime 1a:** Let us first look at the case where the entire post-recuperation phase
is spent in leisure, i.e. \( z = 1 \). This implies that \( \nu \geq 0 \) and \( \lambda \sigma (x) (1 - r) z \geq 0 \). Plugging these
values into the second first order condition (FOC) we see that \( \lambda \geq 1 > 0 \). Furthermore,
from the first FOC we find that \( \lambda \geq 1 \) in turn implies that \( \mu > 0 \) which, due to the
complementary slackness requirement, implies \( x = 1 \). By substituting these values for
recuperation and post-recuperation leisure, \( x \) and \( z \), into equations 3.1 and 3.2, this
procedure gives us the following expression for leisure and labour:

\[
L = s r + (1 - r)
\] (3.4)

\[
\Phi (L) = s r + (1 - r) - L = 0
\]
From these equations we see that total efficient time, $T$, which is equal to efficient labour plus efficient leisure, will be given by:

$$T = L + \Phi(L) = 1 - r(1 - s)$$ (3.5)

This type of regime 1a is illustrated by point 1 in figure 3.1 below.

**Regime 1b:** Next, we look at the case where some of the post-recuperation phase is spent in leisure, i.e. $1 > z > 0$. This implies that $v = 0$ and $[\Phi' / \alpha z] = 0$. Substituting these values into the second FOC we now see that $\lambda = 1 > 0$. From the first FOC we find that $\lambda = 1$ in turn implies that $\mu > 0$ which, due to the complementary slackness requirement, implies $x = 1$. By plugging these values for recuperation and post-recuperation leisure, $x$ and $z$, into equation 1 this procedure once more gives us an expression for leisure:

$$L = sr + (1 - r)z$$

(s $r < L < s r + (1 - r))$$ (3.6)

Hence, the maximum amount of labour obtainable given a certain amount of leisure can once more be expressed as follows:

$$\Phi(L) = sr + (1 - r) - L$$ (3.7)

Differentiating this expression with respect to efficient leisure yields the slope of the efficient time budget constraint: $\Phi'(L) = -1$. As expected, the efficient time budget constraint is negatively sloped in the efficient labour and efficient leisure space. A value of $-1$ means that for the type of individuals that choose to recuperate fully, each additional hour of efficient leisure can be substituted for an hour of efficient labour. Just as in the case of regime (1a) total efficient time will be given by the following expression:

$$T = L + \Phi(L) = 1 - r(1 - s)$$ (3.8)

This type of regime (1b) is illustrated by the line denoted 2 in figure 3.1 below.

The main insight from the previous analysis is that if any leisure is taken in the post-recuperation phase then recuperation is spent entirely in leisure (recuperating). This possibility arises if and only if $L > s r$. The intuition behind this result is clear; if an individual spends any time on leisure he should rather spend it during recuperation than post-recuperation time, because the former choice will at least improve the post-recuperation productivity whereas the latter will not.\(^{20}\) Notice that if the entire recuperation time is spent in leisure and none of the post-recuperation time, i.e. $x = 1$, $z = 0$, then $L = s r$, $H = (1 - r)$, and $\Phi(s r) = 1 - r$.

\(^{20}\) It is easiest to see this by looking at the two extreme cases where all leisure comes from either 1) the post-recuperation phase, or 2) the recuperation phase. Case 1 implies: $\Phi = sr + (1 - z)(1 - r)s$, $L = z(1 - r)s$, and total time, $T$, equals $T = \Phi + S = s < 1 - r + sr$ (because $0 < (1 - r)/(1 - s)$). Case 2 implies: $\Phi = (1 - z)(1 - r)$.
3.3.2 Regime 2: No Post-Recuperation Time Spent in Leisure

The second type of regime occurs when $0 \leq L \leq s\ r$. As we have seen above this implies that no leisure is taken during post-recuperation ($z = 0$). This solution type is characterised by: $0 \leq z \leq 1$, which implies $\mu \geq 0$ and $[\partial \Phi / \partial z] \leq 0$. Substituting these values into the first FOC we then find that $\lambda < l$, which from the second FOC implies $z = 0$. A marginal effect on labour of giving up one unit of leisure, $\lambda$, which is smaller than one implies that less efficient labour can be obtained by trading in efficient leisure. These worsened terms of trade are not due to characteristics specific to the individual, but to the characteristics of the illness. The solution may be usefully split into two scenarios:

**Scenario 1:** In the case where $\lambda > 0$, i.e. where there is some efficient labour to be gained by trading in efficient leisure time, we obtain from the third FOC the following expression for leisure: $L = s\ r\ x < s\ r$. By substituting the ensuing expression for actual recuperation, i.e. $x = L / sr$, into equation (3.2) we derive the expression for efficient labour:

$$
\Phi (L) = s\ r - L + (1 - r) \sigma \left( \frac{L}{sr} \right)
$$

Differentiating this expression with respect to leisure will give us the slope of the efficient time budget constraint:

$L = sr + z (1 - r)$, and $T = 1 - r + s\ r$. 

Source: Own illustration.
The slope is still negative but with an absolute value now smaller than one, which implies that it is less steep than for the full-recuperation regimes. Furthermore, it means that more efficient labour can be gained by working an additional hour during the recuperation time than by recuperating during that hour and thus gaining in post-recuperation productivity, i.e. \( \sigma' \left( \frac{L}{sr} \right) < \frac{sr}{1-r} \).

The second derivative of efficient labour with respect to efficient leisure gives us the curvature of the budget constraint:

\[
\Phi''(L) = \frac{1-r}{sr^2} \sigma'' \left( \frac{L}{sr} \right) < 0
\]

This implies that the budget constraint is concave.

**Scenario 2:** In the case where \( \lambda = 0 \) (i.e. no additional efficient labour can be gained by trading in efficient leisure) leisure is given by the following expression:

\[
L < sr \implies \sigma'(L) = \frac{sr}{1-r},
\]

and that for a certain amount of recuperation time, \( x \), the efficient labour that can be gained by working an additional hour during the recuperation time is equal to that which can be gained by recuperating that hour and thus gaining in post-recuperation productivity, i.e. \( \sigma'(x) = sr/(1-r) \). Knowing that \( x > L/sr \), and that there are decreasing returns to recuperation we can further deduce that \( \sigma' \left( \frac{L}{sr} \right) > \frac{sr}{1-r} \).

We have now seen the general characteristics of the type three regime. In order to really understand what determines the shape of the budget constraint when an amount of efficient leisure is chosen which is less than the point denoted \( sr \) in figure 3.1 (i.e. \( 0 \leq L \leq sr \)), and to derive this line in more detail the maximisation problem will now be simplified. Knowing that when part of recuperation time is spent working there will be no post-recuperation time leisure, let us define our problem as one of maximising consumption/work given that \( z=0 \):

\[
\hat{\Phi} \equiv \max_{0 \leq x \leq 1} sr(1-x) + \sigma(x)(1-r) \tag{3.9}
\]

The maximisation problem gives the following first order condition:
Based on this first order condition we can now distinguish in more detail between 3 different cases which have an impact on the efficient time frontier, each determined by the characteristics of the illness.

**Case 1:** In this case we investigate the conditions under which exact full recuperation, i.e. \(x = 1\), yields the maximum amount of obtainable efficient labour. From the Kuhn-Tucker conditions above we find that \(x = 1\) implies \(\mu \geq 0\) and \(x [\partial \Phi / \partial x] = 0\). This then yields the following condition for case 1:

\[
\sigma'(l) \geq \frac{sr}{l - r} 
\]

Note that \(\sigma'(x)\) is the marginal productivity of an increase in actual time off during recuperation time, measured in terms of the extra productivity in post-recuperation time. Case 1 implies that it always pays to take time off during the recuperation period, even at the point where all recuperation time is spent at home.

In this case the solution to the maximisation problem in (3.9) occurs when \(x = 1\), and \(\Phi = 1 - r\). Here \(\Phi(L) = \Phi = 1 - r\) (\(\forall L, 0 \leq L \leq s\ r\)), which means that the entire recuperation period is spent recuperating and the entire post-recuperation period is spent working (see figure 3.2). Notice that, by concavity, \(\sigma'(l) < \sigma(1) - \sigma(0) = 1 - s\). Using this inequality, together with expression (3.10), we find that a necessary condition for this case to arise is that \(s < 1 - r\). We can conclude that this case will be more likely to occur when the marginal productivity during recuperation is relatively low.

Figure 3.2 depicts the efficient time constraint for case 1. Maximum leisure is obtained when the person stays at home the entire recuperation period (equivalent of \(s\ r\) units of efficient leisure) and thus obtains full efficient leisure in the post-recuperation period, \(1 - r\) (because \(\sigma(x) = 1\)): \(L = s\ r + 1 - r\). Since any unit of work during the recuperation period will be at least offset by the ensuing decrease in post-recuperation productivity at work, the maximum amount of work that can be obtained is the equivalent to the entire post-recuperation period.

Hence, at any point on the leisure axis below \(s\ r\), efficient leisure derived from staying at home during recuperation, the maximum obtainable amount of efficient work will be \(1 - r\).
Case 2: The second case considered is when the maximum efficient labour is obtained by not taking any time off in order to recuperate, i.e. $x = 0$. From the Kuhn-Tucker conditions above we find that $x = 0$ implies $\mu = 0$ and $[\partial \dot{\Phi}/\partial x] \leq 0$. This then gives us the following condition for case 3 to arise:

$$\sigma'(0) \leq \frac{sr}{1-r}$$

This case implies that the marginal productivity in the post-recuperation phase due to an increase in actual recuperation during recuperation time multiplied by the length of the post-recuperation period is smaller than the productivity during the recuperation phase multiplied by its length, even at the point where no recuperation time is spent recuperating. Then the maximisation of (3.7) occurs when $x = 0$, and so $\dot{\Phi} = s$. Notice that by concavity $\sigma'(0) > \sigma(1) - \sigma(0) = 1 - s$.

When combining this finding with the initial condition (i.e. $s r / (1 - r) > \sigma'(0)$), we see that a necessary condition for this case to arise is that $s > 1 - r$ (see figure 3.3). This case will therefore be more likely when the marginal productivity during recuperation is rather high. Maximum efficient labour for all $L$, $0 \leq L \leq s r$, is here given by: $\Phi(L) = s r - L + (1 - r) \sigma(L/s r)$.

Figure 3.3 depicts the time constraint under the case 2. Any amount of efficient leisure during the recuperation time can now be traded against additional work time, with
a positive effect on overall efficient work time. If the person works the entire recuperation period, he or she will get an amount of efficient work time equal to \( s > 1 - r \).

Figure 3.3: Efficient time opportunity frontier in the case where working during the entire recuperation time yields the maximum amount of efficient labour (case 2).

Source: Own illustration.

**Case 3:** We finally analyse the conditions under which less than full recuperation, i.e. \( 0 < x < 1 \), yields the maximum amount of obtainable efficient labour. From the Kuhn-Tucker conditions above we find that \( 0 < x < 1 \) implies \( \mu = 0 \) and \( \frac{\partial \phi}{\partial x} = 0 \). This then yields the following condition for case 3:

\[
\sigma'(x) = \frac{sr}{1-r} \quad (3.12)
\]

Substituting in the highest and lowest possible amounts of actual recuperation, i.e. \( x = 1 \) and \( x = 0 \), and taking into account the fact that we have assumed decreasing returns to recuperation, case 3 then has to fulfil the following condition:

\[
\sigma'(0) > \frac{sr}{1-r} > \sigma'(1) \quad (3.12')
\]

This case implies that there is some amount of actual recuperation between zero and one at which the marginal productivity in the post-recuperation phase of an additional unit of actual recuperation multiplied by the length of the post-recuperation period, is equal to the marginal productivity of an additional time unit spent working during the recuperation phase multiplied by its length. Then there exists an \( x, \hat{x} \), \( 0 < \hat{x} < 1 \), so that
\[ \sigma'(\hat{x}) = sr/(1-r), \] and
\[ \hat{\Phi} = sr(1-\hat{x}) + \sigma(\hat{x})(1-r) > 1-r. \] Then efficient labour is given by the following two expressions:

\[ \Phi(L) = \begin{cases} \hat{\Phi}, 0 \leq L \leq sr \hat{x} \\ sr - L + (1-r)\sigma \left( \frac{L}{sr} \right), sr \hat{x} \leq L \leq sr \end{cases} \]

Figure 3.4 depicts the time constraint under case 3. The maximum amount of leisure available is the same as in case 1: \( L = sr + l-r \). However, a certain amount of efficient leisure during the recuperation time, \( sr(1-\hat{x}) \), can now be traded against additional work time, with a positive effect on overall efficient work time.

Below the point marked \( sr \hat{x} \) on the leisure axis there is no more efficient work to be gained by trade, because any additional work during recuperation (>1-\( \hat{x} \)) will be at least offset by the ensuing decrease in post-recuperation productivity at work.

Figure 3.4: Efficient time opportunity frontier in the case where working part of the recuperation time yields the maximum amount of efficient labour (case 3).

\[ \Phi(L) \]
\[
\begin{array}{c}
\hline
s \\
L-r \\
\hline
0 \ sr \hat{x} \ sr \ sr+1-r
\end{array}
\]

Source: Own illustration.

3.3.3 Interim Summary

The efficient time frontier can be divided into three cases distinguished by their illness characteristics. Each of these cases consists of two main types of regimes. The first regime is characterised by some leisure being taken in the post-recuperation period, as well as all of the recuperation period being spent actually recuperating (i.e. \( x=1, 0 \leq z \leq l \)). This regime will be the same in all three illness-cases, and is illustrated by the straight
line with slope -1 denoted ‘B’ in figure 3.1. The second regime is characterised by no leisure being taken in the post-recuperation period (i.e., \( z=0, 0 \leq x \leq I \)), and is the regime that distinguishes the three illness cases. Case 1 arises when productivity during the recuperation period is relatively low - certainly the efficient labour obtained from working the entire period \( s \) is less than what would be obtained by fully recuperating and working the entire post-recuperation period at full productivity \((1-r)\). In this case everyone will spend the full time recuperating. The cases 2 and 3 both predict that all, or part, of the recuperation time can be traded against additional work time with a positive effect on overall efficient work time, and therefore that some consumers will choose to not fully recuperate. Which of the three cases is the correct description of reality will depend on the type of illness and type of work considered. From here onwards we concentrate on the two pure cases, case 1 and 2, as the third case is really just a hybrid of the two others and does not bring any additional insights into the discussion.

### 3.4 Utility Maximisation

When deriving the efficient time budget constraint we found that there is a tendency for individuals who choose higher amounts of leisure to fully recuperate. The question that now will be explored is therefore what type of individuals tend to spend more time in leisure? The only feature that distinguishes the individuals in this model is the wage rate, so the related question is whether wage has an effect on the amount of efficient leisure chosen.\(^{21}\)

In order to answer the question asked above, i.e. who will choose to recuperate fully, we now turn to maximising utility. Utility, which is assumed to be a function of consumption and efficient leisure, will be maximised subject to the budget constraint, using the insights about maximum obtainable efficient labour time derived in the previous section. The main aim in this part of the analysis is twofold: first, to find the marginal effect of wage upon leisure, and second, to analyse whether a critical wage exists above which individuals choose to recuperate fully.

#### 3.4.1 The Marginal Effect of Wage Upon Leisure

Consumption is split into two types, the necessary minimum requirement, \( m \), (which we assume to be more or less of the same size for all individuals) and the

\(^{21}\) In practice, the level of unearned income may also vary between individuals and we will therefore also briefly discuss the effect of unearned income on leisure choice in appendix 3.D.
consumption over which we have a choice, \( c \). The individual further has unearned income, \( \bar{y} \), and wage earnings, \( w\Phi(L) \), where \( \Phi(L) \) is the maximum efficient labour obtainable for a certain level of efficient leisure, as derived in section 3.3. This gives us the following budget constraint:

\[
c + m \leq \bar{y} + w\Phi(L)
\]

Define \( a = m - \bar{y} \), where \( a \) is the net minimum consumption constraint. We then have the following expression for consumption:

\[
c \leq w\Phi(L) - a
\]  

(3.13)

Utility, a function of consumption and leisure, is maximised subject to the budget constraint:

\[
\max_{0 \leq L \leq 1} U[w\Phi(L) - a, L]
\]  

(3.14)

This yields the familiar first order condition:

\[
U_c w\Phi'(L) + U_L \leq 0
\]  

(3.15)

Following the envelope theorem, the marginal effect of wage upon leisure, i.e. the sign of \( dL(w)/dw \), has the same sign as the following expression:\textsuperscript{22}

\[
\frac{\partial}{\partial w} \left[ U_c (w\Phi(L) - a, L)w\Phi'(L) + U_L (w\Phi(L) - a, L) \right]
\]

Let us assume an additively separable utility function, i.e.: \( U = f(c) + g(L) \).\textsuperscript{23} Then, the sign of \( dL(w)/dw \) will be equal to the sign of the following expression:\textsuperscript{24}

\[
\left\{ 1 - \frac{w\Phi(L)}{(w\Phi(L) - a)} \left( -\frac{c''}{f'} \right) \right\}
\]  

(3.16)

Let us define \( \lambda = (-c''/f') \). The lambda, \( \lambda \), is thus the elasticity of the marginal utility of consumption with respect to consumption, also known as the reciprocal of the elasticity of substitution, \( \sigma \).

The general case of an additively separable utility function has now been briefly introduced. However, the CES (constant elasticity of substitution) utility function gives all the essential insights of the general case and we will therefore conduct the detailed discussion within this framework.\textsuperscript{25} This will in particular give us the advantage of being

---

\textsuperscript{22} See appendix 3.A1 for mathematical explanation.

\textsuperscript{23} Using a utility function that is not additively separable will not significantly change the analysis, see appendix 3.A2.

\textsuperscript{24} For calculations see the appendix 3.A3.

\textsuperscript{25} For a brief discussion of the effect of wages upon leisure with some specific cases of utility functions see appendix 3.A4.
able to analyse in more detail what the value of the elasticity of substitution between leisure and consumption implies for the effect of increased wages upon leisure choice, as well as the conditions under which a critical wage above which individuals choose to fully recuperate exists.

Let us take the following CES utility function:

\[ U(c,L) = \left( \frac{\alpha}{1-\lambda} \log(c) + \frac{1-\alpha}{1-\lambda} \log(L) \right)^{-\lambda}, \lambda \geq 0, \lambda \neq 1 \]

where \( 0 < \alpha < 1 \), and \( \sigma = 1/\lambda \) is the elasticity of substitution. The Cobb Douglas case is where \( \sigma = \lambda = 1 \). Maximising this utility function with respect to the budget constraint can therefore be expressed as:

\[ \max_{c,L} U = \frac{\alpha}{1-\lambda} [w\Phi(L) - \alpha]^{-\lambda} + \frac{1-\alpha}{1-\lambda} L^{-\lambda} \]

The first order condition is:

\[ \alpha w\Phi'(L) [w\Phi(L) - \alpha]^{-\lambda} + (1-\alpha) L^{-\lambda} = 0 \]

Then, following the envelope theorem, the sign of \( \frac{dL(w)}{dw} \) will be equal to the sign of:

\[ \frac{1}{1-\lambda} \left( \frac{w\Phi(L)}{w\Phi(L) - \alpha} \right)^\lambda \]

Define \( \mu = \frac{w\Phi(L)}{w\Phi(L) - \alpha} \), where \( \mu \) is the ratio of overall consumption to choice consumption, the latter being given by total income net of the net minimum consumption requirement. Whether the expression \( \lambda \mu - 1 \) overall has a positive or negative sign thus depends on whether \( \lambda \mu = \mu/\sigma > 1 \) or not.

The ratio of overall consumption to choice consumption is always larger or equal to one, \( \mu \geq 1 \). When the wage rate goes to infinity \( \mu \) goes to one (i.e. \( \mu \to 1 \) if \( w \to \infty \)), and when the wage rate goes to the existence wage (\( w \to w_{\text{min}} = \alpha/\Phi(L) \)) \( \mu \) goes to infinity (i.e. \( \mu \to \infty \) if \( w \to w_{\text{min}} \)).

---

26 The textbook CES-function is often defined as: \( U = [a_1 x_1^\rho + a_2 x_2^\rho]^1/\rho \). Since preferences are invariant with respect to monotonic transforms of utility, we can just as well use equation (3.17). Note: \( \rho = 1-\lambda \). The term \( 1-\lambda \) in the denominator is needed to ensure that the utility function is strictly increasing in \( c \) and \( L \) (i.e. that the marginal utilities of consumption and leisure are positive even in the case where \( \lambda > 1 \)).
The higher \( \lambda \), and thus the lower the elasticity of substitution between consumption and leisure, the more rapid the proportionate decline in the marginal utility of consumption in response to increases in \( c \). Hence, the less willing individuals are to accept deviations from a uniform pattern of consumption due to e.g. sickness. Since we are looking at one period only, these changes in consumption can be thought of as being across individuals or as hypothetical alternative scenarios for the individual. As long as the utility function is concave in \( c \), an accepted and realistic assumption to make (risk aversion with respect to consumption), \( \lambda > 0 \). Estimation of the elasticity of substitution, \( \sigma \), has tended to find a value smaller than one, and this implies \( \lambda > 1 \). As long as the elasticity of substitution is smaller than one, leisure will increase in wages since the ratio of total consumption to choice consumption was found to be larger than one. Furthermore, the closer the wage rate comes to the existence wage, the higher will be the elasticity of substitution that still allows leisure to be increasing in wages.

Based on econometric work on micro data from Norway by Aaberge et al. (1995) the elasticity of substitution between consumption and leisure was found to be 0.25. Pencavel (1986) reviewed a number of empirical studies with elasticity estimates for men in the United States and Britain. The substitution effect was found to fall in the range of 0.1 to 0.2 in Britain, and to be about 0.1 for the United States. Blundell and MaCurdy (1998) similarly summarised the results from a number of empirical studies and found that the elasticity from the majority of the studies was positive and smaller than 1.

Even if \( \lambda \) were to be less than one, for example \( \lambda = 0.7 \), leisure would be increasing in wages as long as \( \mu > \lambda / 0.7 \), which implies that the ratio of net minimum consumption requirement to wage income, \( a / w \Phi (L) \), must be larger than 0.3. In other words, if the elasticity of the marginal utility of consumption with respect to consumption were 0.7 (i.e. if the elasticity of substitution were 1.4), then the minimum consumption requirement on wages has to be at least 30% of maximum wage earnings (for any given amount of leisure) in order for the marginal effect of wage on leisure to be positive.

### 3.4.2 The Existence of a Critical Wage Above Which People Choose to Fully Recuperate

We will now turn to the two types of regimes discussed in section 3.3 and analyse the conditions that determine the existence and size of a critical wage above which people choose to fully recuperate. Under regime 1 people choose to fully recuperate and even
take some post-recuperation leisure time, whereas under regime 2 people choose to work part of the recuperation time or exactly recuperate (regime 1a discussed previously, where the whole period is spent in leisure and none in labour, is ruled out by our assumption of a positive net minimum consumption requirement).

Regime 1: In order to analyse the amount of leisure chosen by those who at least recuperate fully we can use equation (3.5) or (3.8) to substitute for $\Phi(t)$ and $\Phi'(t)$. From equation (3.18) the expression for leisure is derived:

$$L = \frac{w T - a}{\gamma w^\sigma + w}$$  \hspace{1cm} (3.19)

where $\gamma = \left(\frac{\alpha}{1 - \alpha}\right)^\sigma$ and $a > 0$ is the net minimum consumption. We know from the analysis in section 3.3 that the wage group that will choose to at least fully recuperate fulfils the condition of $L \geq s r$. Hence, we focus on the case where there is a critical wage, $\bar{w}$, for which $L(\bar{w}) = L = sr$, and $L(w) > L$ for all wages above the critical wage ($\forall w > \bar{w}$). $L, 0 < L < T$, is the critical value of leisure above which people choose to fully recuperate.

Differentiating equation 3.19 with respect to wage we obtain:

$$\frac{dL}{dw} = \frac{(1 - \sigma) w T + a (I + \gamma w^{-(1 - \sigma)})}{[\gamma w^\sigma + w]^2} > 0, \quad \text{if } \sigma \leq 1$$  \hspace{1cm} (3.20)

Leisure demand is hence strictly increasing in wage as long as we assume $\sigma \leq 1$.

From equation (3.19) we observe that if the income obtained by working the entire period exactly equals the net minimum consumption requirement, then no leisure is taken (i.e. $L = 0$ when $w T = a$), whereas when wages tend towards infinity leisure time tends towards total time (i.e. $L \to T$ as $w \to \infty$). Hence there always exists a unique $\bar{w}$ for which $L(\bar{w}) = L$.

---

27 The budget constraint given in equation (3.13) ($c = w \Phi(t) - a$) implies that total consumption (consumption, $c$, plus the minimum consumption requirement, $m$) in the type 1 regime must be smaller or equal to unearned income, i.e. $c + m \leq \bar{y}$. This again will imply that the net minimum consumption requirement is equal to, or less than 0 ($\alpha \leq 0$). Given a fixed level of required minimum consumption, this type of regime is therefore only possible for high levels of unearned income, $\bar{y}$. Certainly, the amount of unearned income has to be equal to or larger than the required minimum consumption, i.e. $\bar{y} \geq m$. An interesting extension to the model would be to make unearned income a function of income saved from previous periods. In that case, and assuming that the current wage rate is a function of previous wage rate, the wage rate would not be irrelevant for this type of regime.
In the Cobb-Douglas case, where \( \sigma = 1 \), we see that once again no leisure is taken when earning the existence wage (i.e. \( L = 0 \) when \( w T = a \)), whereas when wages tend towards infinity leisure time now tends towards an amount less than total time (i.e. \( L \rightarrow T/(1 + \gamma) \) as \( w \rightarrow \infty \)). In order to ensure that there is a critical wage \( \bar{w} \) for which \( L \leq T/(1 + \gamma) \), it is in this case necessary to assume that \( \bar{L} \leq T/(1 + \gamma) \).

Analytical expressions for the critical wage could be derived in at least two cases, the first being the Cobb-Douglas case where \( \sigma = 1 \), and the second is where \( \sigma = 0.5 \), and the results can be found in appendix 3.C.

In order to give an idea as to what happens to the critical wage for various values of the elasticity of substitution, three numerical examples will now be analysed. In each case we assume that \( a = 1 \). Example 1 is defined by \( \alpha = 0.5 \), \( s = 0.5 \), and \( r = 0.3 \), example 2 by \( \alpha = 0.2 \), \( s = 0.5 \), and \( r = 0.3 \), and example 3 by \( \alpha = 0.5 \), \( s = 0.5 \), and \( r = 0.8 \). The parameter values are substituted into equation (3.19), and different wage rates are simulated for the different values of \( \sigma \) in order to find the critical wage, \( \bar{w} \), that equalises the two sides.

When comparing examples 1 and 2 in table 3.1 we observe that the lower \( \sigma \) the lower the critical wage above which people choose to fully recuperate if \( \alpha \geq 1 - \alpha \), whereas in the opposite case the opposite occurs. Furthermore, when comparing examples 1 and 3 it is clear that a longer duration of the recuperation period requires a higher critical wage rate (i.e. the wage rate that separates those who fully recuperate from those who do not).

A wage of 1.43 allows the individual to exactly survive in examples 1 and 2 when working the entire period, given that the minimum consumption requirement on wage income is assumed to be 1 in this example. Hence, with an elasticity of substitution of e.g. 0.8 individuals with a wage rate more than 1.24 times (1.77/1.43) as high as the minimum (net) consumption requirement will choose enough leisure time to at least recuperate fully in the case of example 1.

**Regime 2:** We can calculate the range of wages for which individuals would choose to exactly recuperate for the first illness case, i.e. \( \sigma'(1) = \beta (1 - s L^g) \geq s r/(1 - r) \). It was shown that for this sickness case the total effect on efficient working time from working during illness would not be positive, and hence no one would work during recuperation. The existence wage is then given by the net minimum consumption requirement divided by the maximum obtainable efficient labour time:

28 See the footnote under table 3.1 for the calculation of this existence wage.
\[ w_{\text{min}} = \frac{a}{(1 - r)} \leq w \]  

(3.21)

Hence, at any wage between the existence wage, \( w_{\text{min}} \), and the critical wage, \( \bar{w} \), the individuals will exactly recuperate. In examples 1 and 2 in table 3.1 the survival wage for case 1 has been illustrated.

In the second illness case, where \( \frac{sr}{1 - r} \geq \sigma'(0) = \beta^s(\rho^{2-1}b(1 - s^{1/b})) \), we found the following expression for maximum working time for given amounts of leisure:

\[ \Phi(L) = sr - L + (1 - r)\sigma(L) \left( \frac{L}{sr} \right) \quad \text{for} \quad 0 \leq L \leq sr \]  

(3.22)

Substituting for \( x=0 \) in equation (3.22) gives us the following existence wage:

\[ w_{\text{min}} = \frac{a}{s} \leq w \]  

(3.23)

Hence, at any wage between the existence wage and the critical wage individuals will choose an amount of efficient leisure between \( 0 \leq L \leq sr \). The exact amount can only be derived numerically after having assumed different parameter values. In example 3 table 3.1 the survival wage for case 2 has been illustrated.

### 3.5 Summary

This chapter has considered the choice an individual faces on how to spend her sickness time, given that this choice affects the post-recuperation productivity of labour, as well as the utility derived from leisure (productivity of leisure) during that period. The analysis finds that if any leisure is taken in the post-recuperation phase then recuperation is spent entirely in leisure (recuperating). The reason for this is that recuperating during the recuperation phase has the additional advantage of increasing the productivity in both work and leisure in the post-recuperation phase. Hence, when the individual chooses to take out the equivalent of the recuperation time or less in leisure she will only take it during the recuperation time. How much of this time she chooses to work will, however, depend on the productivity-decreasing effect of the illness, as well as on the wage level, and the minimum consumption requirement.

An implication of the above finding is that the more leisure an individual takes, the more likely she is to stay at home during recuperation. Given a CES utility function and an empirically supported elasticity of substitution between leisure and consumption, and assuming that the minimum consumption constraint on wage income net of unearned income is non-negative, there will be a tendency for an increase in the wage to increase the demand for leisure. In this case the analysis suggests that people with low wages...
(below a certain level determined by preferences and the budget constraint) will be working during sickness time, which in turn will imply a lower productivity in their post-recuperation phase. The adverse health effect from a given illness is therefore influenced by the choice made on how to spend ones time, and will be larger for low-wage individuals than higher-wage individuals.

The income and welfare implications of this finding, as well as its implication for traditional illness measurements are discussed in chapter 4.
Table 3.1: Critical wage (i.e. a wage rate above this wage yields post-recuperation leisure) for different values of $\sigma$ ($a=1$).

<table>
<thead>
<tr>
<th>$\sigma$</th>
<th>Example 1: $a=0.5$, $s=0.5$, $r=0.3$</th>
<th>Example 2: $a=0.2$, $s=0.5$, $r=0.3$</th>
<th>Example 3: $a=0.5$, $s=0.5$, $r=0.8$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$w_{\text{min}}$</td>
<td>$\bar{W}$</td>
<td>$w_{\text{min}}$</td>
</tr>
<tr>
<td>0.10</td>
<td>1.43</td>
<td>1.65</td>
<td>1.43</td>
</tr>
<tr>
<td>0.20</td>
<td>1.43</td>
<td>1.67</td>
<td>1.43</td>
</tr>
<tr>
<td>0.30</td>
<td>1.43</td>
<td>1.68</td>
<td>1.43</td>
</tr>
<tr>
<td>0.40</td>
<td>1.43</td>
<td>1.69</td>
<td>1.43</td>
</tr>
<tr>
<td>0.50</td>
<td>1.43</td>
<td>1.71</td>
<td>1.43</td>
</tr>
<tr>
<td>0.60</td>
<td>1.43</td>
<td>1.73</td>
<td>1.43</td>
</tr>
<tr>
<td>0.70</td>
<td>1.43</td>
<td>1.75</td>
<td>1.43</td>
</tr>
<tr>
<td>0.80</td>
<td>1.43</td>
<td>1.77</td>
<td>1.43</td>
</tr>
<tr>
<td>0.90</td>
<td>1.43</td>
<td>1.79</td>
<td>1.43</td>
</tr>
<tr>
<td>1.00</td>
<td>1.43</td>
<td>1.82</td>
<td>1.43</td>
</tr>
</tbody>
</table>

Source: Own calculations.

* A wage rate of one ($w=1$) allows a healthy individual to exactly survive when working the entire period (normalised to one), given that the minimum consumption requirement on wage income is assumed to be 1 ($\sigma=1$). In example 1, the parameter values fulfil a necessary condition for being a sickness case 1: $s < 1-r$. With a case 1, no additional efficient time can be gained by working during recuperation, hence, maximum efficient working time is the exact post-recuperation time. Survival wage in this case is therefore $w_{\text{min}} = a/(1-r)$ ($=1/0.7=1.43$). In example 3, the parameter values fulfil a necessary condition for being a case 3: $s > 1-r$. With a case 3, additional efficient time can be gained by working during recuperation, hence, maximum efficient working time is total real time ($=1$) multiplied by the productivity during sickness ($s$). Survival wage in this case is therefore $w_{\text{min}} = a/s$ ($=1/0.5=2.00$).
Chapter 4: Measuring the Adverse Effects from Illness

4.1 Introduction

In the previous chapter we saw that optimising individuals who derive some leisure out of staying at home during the recuperation time from an illness will only choose leisure in the post-recuperation time if the entire recuperation time was spent at home. This implied in particular that those who tend to recover fully choose to have relatively much leisure time. Given a net minimum consumption requirement, it was further found that it is individuals with a higher wage rate who tend to choose much leisure time and therefore to recover fully before returning to work.

In the present chapter different measurements of the adverse health effect from illness are discussed, in particular absence from work and the amount of efficient time, income, and welfare loss are analysed. The losses in different wage groups are compared, and the differences in predictions between the exogenous (traditional) and the endogenous recuperation model are presented. The losses from illness are as far as possible derived analytically for different wage groups, and two numerical examples corresponding to the two different illness cases discussed in chapter 3 are presented.

Section 4.2 discusses the efficient time loss from illness for different income groups and compares it to absence from work as a measure of the adverse effect of illness, whereas sections 4.3 and 4.4 analyse the income and welfare losses, respectively. The two examples are presented in section 4.5, and, finally, the findings are summarised in section 4.6.

4.2 Efficient Time Loss (ETL) from Illness and Absence from Work (WA)

One commonly used statistical measurement of illness has been absence from work due to illness, also known as work loss days (e.g. Zuidema and Nentjes (1996), Ostro (1983), Hansen and Selte (2000)). Work-absence will capture the sickness of those who choose to take time off to recuperate, as well as of those who are forced to recuperate because of the severity of the illness. If there is a tendency for low-wage individuals, in the case of less severe illnesses, to work although being ill in order to earn enough to sustain a minimum level of consumption then the health effect for these wage-groups would not appear, or at least be underestimated, using work-absence as adverse effect measurement. If working during the illness period increases the severity and extends the
duration of the illness, then these low-wage individuals may in fact experience a larger adverse health effect than their richer counterparts. In the analysis below we define work-absence (WA) as absence from work during which the individual is sick, and absence is measured as the amount of labour supplied during the illness subtracted from the amount the person would have been working during that period had he or she been healthy. The fact that the individual may be compensating by working more when recuperated does not influence this measure.

Efficient time loss (ETL) can be considered one possible measurement of the adverse health effect of illness that avoids the problem of wage-group bias, as it measures the amount of time a person is ill weighted by the severity of the illness (i.e. time measured in productivity units). This implies that when an individual is less than fully productive (here: productive in terms of producing consumption/income from labour time and/or producing utility from leisure time) a “real” hour is worth less than one when measured in efficiency units, if we normalise healthy productivity to unity. In the model presented in chapter 3 the individuals only used their time for work or leisure, and efficient time is thus the sum of the efficient time spent in these two activities. What this measurement does not take into account is that the efficient time spent in leisure and labour may not be worth equally much to the individual, and that under certain circumstances an individual with less efficient time during an illness may be better off than an individual with more efficient time. ETL from illness is essentially calculated by subtracting the efficient time of an individual when ill from the efficient time when healthy.

Let us assume that all individuals have identical CES utility-functions and that both leisure and consumption are normal goods. With the efficient time budget constraints under healthy conditions and under sickness (developed in the previous chapter) we can now compare the ETL from sickness for different income groups. In order to fulfil the minimum consumption requirement, which is assumed to be the same for all individuals, the individuals with lower wages are required to work more efficient hours than those with higher wages. Thus, a low wage shifts the minimum efficient working-time-constraint (denoted $\bar{a}/w$ in figure 4.2) for the poor upwards. However, whether working while sick yields additional efficient labour time, or whether it is at least offset by the decrease in post-recuperation productivity (i.e. net increase in efficient labour time is smaller or equal to zero), is determined by the illness characteristics. The two illness cases are discussed separately below, but first we analyse the case of the higher-wage individuals for whom the illness characteristics are irrelevant in the sense
that they will always choose to stay home and recuperate fully from their illness before returning to work. The exogenous recuperation case is discussed in conjunction with the higher-wage case, since they are essentially the same.

4.2.1 ETL and WA in the Full-Recuperation and the Exogenous-Recuperation Cases

In the endogenous recuperation (ENDOR) model higher-wage individuals choose not to work while being sick (regime 1), whereas for all wage groups in the exogenous recuperation (EXOR) model sickness time is time lost from \textit{any} activity. Note that we will assume the time loss to be equal across wage groups. In the health capital literature,\textsuperscript{29} the time lost to illness is endogenous and a function of investment in health capital, as well as the rate of depreciation. Several of these studies have pointed out that the time lost is likely to vary with individual's wage rates (e.g. Grossman (2000), Cropper (1977)).

In both the ENDOR and EXOR models the individuals will be absent from work during their illness, and the amount of absence measured will depend on how much they would have worked during that period had they been healthy. Below we will analyse geographically and numerically the efficient time loss due to illness in both the regime 1 ENDOR model and the EXOR model, and compare the findings to those obtained when using work-absence.

In figure 4.1 we observe that at full "healthy" time the high-wage individuals in the ENDOR model and individuals in the EXOR model will choose point A - the point of tangency between the healthy budget constraint and the healthy indifference curve, $U^H$. When there is some illness during the period, the former individuals move to the lower efficient time constraint (the endogenous sick budget constraint), a lower indifference curve, $U^S$, and thus to point $B$. Since these individuals always choose to recuperate fully, the fact that the terms for trading in leisure for labour worsen when the individual decides to work during the recuperation time is irrelevant. The worsened terms of trade are captured in figure 4.1 by the fact that the upper part of the endogenous sick budget constraint lies somewhere within the coloured triangle.

The conventional way of modelling illness as time lost, i.e. time that can not be used in any other activities, is captured by a budget line which is parallel to the healthy one but shifted inwards. Hence, there is very little analytical difference between the endogenous recuperation model in the case of those with relatively high wages, and the exogenous recuperation model. In the latter, however, the time spent home recuperating will not provide any leisure to the individual and the exogenous efficient time budget
constraint will therefore lie to the left of the linear part of the *endogenous* sick budget constraint in the diagram (total efficient time available would be \(1-r\)). The individual will end up at point \(C\) in figure 4.1, the tangency between the exogenous sick budget constraint and their indifference curve, \(U_x\).

When it comes to measuring analytically the efficient time loss experienced by higher-wage individuals due to sickness, we have seen that their utility functions are tangent to the "sick" time budget constraint on the linear part of it, i.e. where the marginal rate of substitution is equal to one as in the healthy scenario. Given that we are in the case where the recuperation time is taken out in leisure, the amount of efficient time is independent of the amount of leisure taken (in addition to this minimum). The efficient time loss (ETL) is calculated by subtracting the efficient time under illness (based on equations (3.5) or (3.8) in section 3.3) from the efficient time under healthy conditions (which was normalised to one). The subscript 1 indicates that we are looking specifically at the high-wage individuals of regime 1:

\[
ETL_1 = [1 - (1 - r + sr)] = r (1 - s), \quad \frac{\partial ETL}{\partial r} = 1 - s > 0, \quad \frac{\partial ETL}{\partial s} = -r < 0 \tag{4.1}
\]

In words, these individuals loose the recuperation time, \(r\), less what it is worth in terms of efficient leisure time, \(sr\), where \(s\) is the productivity during illness. Hence, the longer the full recuperation period, and the lower the productivity during the recuperation period, the larger the efficient time loss.

In the exogenous recuperation model the efficient time loss will be larger than that experienced in the endogenous-recuperation-model by a factor of \(1/(1-s)\):

\[
ETL_X = [1 - (1 - r^-)] = r
\]

The subscript \(X\) indicates that we are looking at the exogenous recuperation model. The loss is larger because recuperation time is not worth anything in terms of leisure time in this model. Hence, the longer the full recuperation period the larger the efficient time loss.

Using work-absence as a measure of illness we find that all wage groups in the EXOR model, as well as the high-wage individuals in the ENDOR model are absent from work during their illness. The period when they are ill would have been like any other time period under healthy conditions, hence it would have been used partly for leisure, partly for choice consumption, and partly to fulfil the net minimum consumption requirement during that sub-period. We can obtain the analytical expression for labour in the Cobb-Douglas case, a special case of CES utility where the elasticity of substitution is

---

29 Refer to chapter 2, section 2.2.1 for a review of this literature.
1, by substituting the expression for leisure derived in (3.19) into the expression for labour (3.7) and solving it for $\sigma = l$. The resulting expression gives the fraction of any time period that the individual would have worked when healthy, and by multiplying it by the amount of recuperation yields an expression for $WA$. Absence from work due to illness will therefore, both in the exogenous case and the regime 1 endogenous case, be given by the following expression:

$$WA_i = r \left( \alpha \left( l - \frac{a}{w} \right) + \frac{a}{w} \right)$$

with $\frac{\partial WA}{\partial w} < 0 \quad (4.2)$

The $\alpha$ denotes the weight given to choice consumption in the utility function, and $a$ is the net minimum consumption requirement. The amount of absence from work will decrease with increasing wages due to the fact that the amount of leisure time chosen increases with wages, and corollary with this finding the time worked decreases with wages and therefore the working time lost due to illness.

ETL is hence the same for all wage groups in the EXOR model and for all those individuals in the ENDOR model who choose to recuperate fully, however, it is smaller for the latter than the former. In both cases they will stay away from work during their illness, so that work-absence would pick up, at least partly, the adverse health effect. The amount of work-absence at any particular wage-rate will be the same in both models, and it will be decreasing with higher wages.

Figure 4.1: Efficient time loss in the full recuperation and the exogenous recuperation cases.

Source: Own illustration.
4.2.2 Illness Case 1

This section explores the efficient time losses experienced by the low-wage individuals in illness case 1. Due to the illness characteristics, working during the full recuperation time in illness case 1 does not yield additional efficient labour time. Let us remind the reader that the time over which we have preferences, the choice time, is that which comes in addition to the time needed to fulfil the net minimum consumption requirement (the latter is denoted \( a/w \) in figure 4.2 below). Hence, the lower the wage rate the further the individual will be pushed towards the horizontal part of the efficient time budget constraint in figure 4.2, implying that the individual will not be able to work while being sick. Hence, similarly to the high-wage individuals, the low-wage individuals in illness case 1 will be absent from work during their entire illness, and the amount of absence measured will depend on how much they would have worked during that period had they been healthy. The findings of the two illness-measurements, i.e. efficient time loss and absence from work, are quite different, with absence from work performing distinctly better than efficient time loss in this illness case.

When there is no sickness the low-wage individual in figure 4.2 will locate at point \( C \), which is the tangency of the healthy budget constraint and the healthy indifference curve of the “poor”, \( U_p^h \), whereas with sickness point \( D \) will be chosen which lies on the kink. At the kink, the slope of the sick indifference curve of the poor, \( U_p^s \), is smaller (in absolute numbers) than the slope of the budget constraint when full recuperation is chosen, and larger than the slope of the budget constraint when working part of the recuperation period. We see that the low-wage individuals therefore end up loosing the same amount of efficient time as do the higher-wage individuals whose indifference curves are denoted by the subscript \( R \), for “rich”. The analytical expression for the ETL was given in equation (4.1).

Using efficient time loss as an illness measurement is in this case quite misleading. This measurement would make us conclude that illness has the same effect on all wage groups, just as in the EXOR model (although in the latter the estimated effect would be larger), when in fact the low-income individuals would have wanted to trade more of their leisure for work but are forced to recuperate fully. The opportunity to earn additional income by working during the recuperation period, illustrated in figure 4.2 by the coloured triangle, has been removed for the low-wage individuals in case 1 and these are therefore worse off than their richer counterparts.
As was the case for the higher-wage individuals, the low-wage individual in illness case 1 will stay at home during the entire sickness-experience and absence from work will again be given by the expression (4.2) in section 4.2.1. Because of their low wages the individuals discussed here would have been working more had they been healthy, and measured work-absence is therefore larger.

Figure 4. 2: Efficient time loss in illness case 1.

Source: own illustration.

4.2.3 Illness Case 2

Finally we turn to analysing the efficient time losses experienced by the low-wage individuals in illness case 2. In this illness case, trade-off between efficient leisure and efficient work is possible during the full recuperation period. From figure 4.3 we see that the lower the wage rate the further the individual will be pushed towards the concave part of the efficient time budget constraint. This implies that the efficient time over which the individual has a choice decreases, and that the individual may have to work while being sick. Absence from work will therefore yield lower estimates of the adverse health effect from a given illness for the lower-wage individuals than for those with a higher-wage. The findings of the two illness-measurements, i.e. ETL and WA, are once more quite different with the former performing much better than the latter in this case.
When there is no sickness the low-wage individuals will once again be at point C in figure 4.3, whereas with sickness point D will be chosen (tangency of 'sick' budget constraint and the indifference curve, \( U_p \)). It is clear from figure 4.3 that any individual who ends up choosing a point on the concave part of the "sick" budget constraint looses more efficient time than the individual who chooses a point on the part of the "sick" budget constraint which is parallel to the "healthy" one.

For low-wage individuals the slope of the sickness budget constraint at the point where the indifference curve is tangent to it is flatter than the slope of the healthy budget constraint. This implies that in order to obtain one additional unit of efficient leisure the individual will only be willing to give up less than one unit of efficient labour. Therefore, the efficient time loss from sickness is larger for the low-wage than for the higher-wage individuals, i.e.

\[
\text{Loss}_{2,2} > r (1 - s).
\]

In particular, we get the following expressions for efficient time loss (based on the expression of case 2 and 3 from chapter 3 section 3.3):

\[
ETL_{2,2} = [l - ((1-r)\sigma(x) + s r)] = r(l-s) + (l-r)/(l-\sigma) > r(l-s) \quad (4.3)
\]

We can therefore conclude that in illness case 2 low-wage individuals loose more efficient time from illness than do higher-wage individuals. The reason for the larger loss is that due to having worked during the full recuperation time, the productivity in the post-recuperation time has suffered. The size of the difference is determined by the expression \((l-r)/(l-\sigma)\), the lost productivity during the post-recuperation period due to having worked during the recuperation period. It is therefore clear that the longer the recuperation time, and thus the shorter the post-recuperation time, the smaller the difference in ETL between the low-wage and the higher-wage individuals. Similarly, the larger the post-recuperation productivity, other things equal, the smaller the difference in efficient time loss between the poor and the non-poor.

The fact that the low-wage individuals in the illness case 2 face worsened terms of trade for their leisure time because they choose to work while being sick is to some extent captured when using ETL as an illness measurement. The opportunity to earn additional income by working during the recuperation period has been restricted, but not removed, for the low-wage individuals in illness case 2 and these are therefore worse off than their richer counterparts, something also suggested by the higher ETL they experience.

Because the low-wage individuals in this illness case will work while being ill, estimates of the adverse health effect based on work-absence will under-represent the
adverse health effect from illness and yield lower estimates for the lower-wage individuals than for those with a higher-wage. They will also be less absent from work than if they had experienced a more serious illness (illness case 1). The lower the wage below the critical wage, the less days of work-absence will they have until the existence wage is reached where the individual works the entire period in order to survive. Below a certain wage they will end up working more when ill than when healthy. Work-absence will in this case be given by the following expression:

\[ WA_{2,2} = r \left( \alpha \left( 1 - \frac{\alpha}{w} \right) + \frac{\alpha}{w} (1 - x) \right) \]  

\[ (4.2') \]

where \( x \) is the amount of actual recuperation chosen during the full recuperation period (i.e. the part of the recuperation period not spent working). The subscript "2, 2" denotes that we are looking at the low-wage individuals (regime 2) in illness case 2. By setting expression (4.2') equal to zero we can derive the wage rate at which no WA will be registered:

\[ w = \frac{\alpha(1 - \alpha)}{(1 - \alpha - x)} \]

An individual earning a wage rate lower than this will actually work more during the recuperation period than he/she would have done had he/she not been ill during that period.

Figure 4.3: Efficient time loss in illness case 2.

Source: own illustration.
We have seen in this section that the traditional way of modelling illness as time lost, i.e. recuperation as exogenous, misrepresents the adverse effect of illness if illness is in fact a time of choice. In particular, when measuring the effects in this traditional model by the efficient time loss all wage-groups experience the same losses (see figure 4.5), whereas by absence from work the lower the wage the higher the absence (similar to illness case 1 in figure 4.4 below). In the case of a not so severe illness (illness case 2), however, low-wage individuals may choose to work part of their recuperation time and it is shown that this may extend the time period when their productivity is low. The efficient time loss would hence be larger for this wage-group than for the higher-wage ones, as illustrated in figure 4.5, whereas the recorded work-absence would be lower, something which is apparent in figure 4.4. Work-absence therefore performs particularly badly in illness case 2. In the case of more severe illnesses (case 1) where there is nothing to be gained by working while being sick, the efficient-time-loss would record equal illness effects for all wage groups (see figure 4.5) whereas work-absence would increase with lower wages (shown in figure 4.4). Hence, the performance of the two illness measurements has in this case reversed, with efficient time loss being distinctly worse.

Therefore, work-absence would most severely misrepresent the adverse health effects for the low-wage individuals in illness case 2, and efficient time loss would most severely misrepresent the illness effects in illness case 1. The loss experienced due to illness may be better captured by calculating income loss because this latter picks up the removal of the opportunity to earn income in illness case 1, and the constraint on this opportunity in illness case 2. Income losses are the subject of the next section of this chapter.
Figure 4. 4: Work-absence at various wage rates for illness cases 1 and 2.

Figure 4. 5: Efficient time loss at various wage rates for illness cases 1 and 2, and for the EXOR-model.

Note: The figures above are based on the numerical examples given in section 4.5 of this chapter. We have assumed a duration of the full recuperation, r, of 0.3 (out of full time which is unity), a weight on consumption in the utility function, α, of 0.5, and a net minimum consumption requirement, α, of 1. In addition we have assumed the productivity during illness, s, in the case 1 to be 0.5 and in the case 2 to be 0.8.
4.3 Income Loss from Illness

Another commonly used measurement of adverse health effects looks at the change in discounted life-time earnings and personal and societal expenditure on medicines and medical treatment. The latter are not considered in the present model, and since the model is essentially a one period model this Cost-Of-Illness (COI) approach therefore boils down to measuring the income loss caused by illness.

Income loss as a measurement of the effect of illness differs from work-absence and efficient time loss in two important aspects. Since workers are paid according to productivity, income loss due to illness would record the adverse effect on both those individuals who stay at home and of those who go to work, contrary to work-absence. Furthermore, whereas the measurement of efficient time loss implicitly counts leisure time and labour time as equally valuable, income loss looks only and explicitly at efficient work time loss.

When used to calculate the monetary value of health-impacts, the COI approach has been criticised for ignoring the direct welfare aspect of being healthy. It has, however, the advantage of being based on more accessible information than any measurement taking individual preferences into account, and is a measurement that is still more likely to be understood and be taken seriously by policy-makers than the latter. More importantly, income loss from illness is of interest in its own right. Policy-makers are increasingly looking at the effects of different policy-measures on income-distribution, and the income loss from illness of different wage-groups is therefore of interest. In this section the income loss of different individuals with different wages will therefore be derived and compared to the predictions of the exogenous-recuperation model.

In order to derive the income losses, we will need to derive the amount of labour and leisure chosen for different wage groups. We are still assuming a CES-function and, in addition, an assumption on the functional form of post-recuperation productivity will be necessary for the cases where the individual chooses an amount of actual recuperation time less than full recuperation. With these specifications the optimal amount of leisure and work can be calculated for different income groups and for the healthy and sick case. The absolute income loss from illness is then calculated by multiplying the wage rate with the amount of efficient hours worked in the healthy and sick case, and then subtracting the latter from the former. Relative income loss is then the ratio of the absolute income loss to the income earned in either the healthy or the sick case. We have chosen to use the healthy income for this purpose.
4.3.1 Income Loss in the Full Recuperation and the Exogenous Recuperation Cases

In order to compare the income losses due to illness of different wage-rate groups, we first need to derive individuals’ optimal amount of leisure chosen when they are healthy the entire period. This efficient leisure time will then be subtracted from total efficient time available in order to give optimal efficient work time, which multiplied by the wage rate gives healthy income.

We remind the reader of the assumed preference function, the CES utility function:

$$U(c,L) = \begin{cases} \frac{\alpha}{1-\lambda}[C]^{\frac{\lambda}{1-\lambda}} + \frac{1-\alpha}{1-\lambda}[L]^{\frac{\lambda}{1-\lambda}}, \lambda \geq 0, \lambda \neq 1 \\ \alpha \log(c) + (1-\alpha)\log(L), \lambda = 1 \end{cases}$$

where $c$ is consumption and $L$ is leisure, and where $0 < \alpha < 1$, and $\sigma = \lambda = 1$. The Cobb-Douglas case is where $\sigma = \lambda = 1$.

An individual will either work or spend time in leisure:

$$H = T(L) - L$$

where $H$ is efficient labour, $L$ is efficient leisure, and $T$ is total efficient time. Let us for later convenience also define “choice” labour, $H_C$, and total “choice” time, $T_C$, as the amount of labour and total time over and above what is required to fulfil the net minimum consumption requirement (i.e. $H_C = T(L) - L - a / w$, $T_C = T - a / w$, where $a$ is the net minimum consumption requirement, and when divided by the wage rate, $w$, gives the amount of labour hours required to fulfil the requirement). Total time was assumed equal to one in the healthy case ($T^H = 1$), and in the regime 1 (chapter 3, section 3.3) it was found to be given by the following expression: $T^S = 1 - r(1-s)$. Superscripts $H$ and $S$ indicate the healthy and sick scenario, respectively, and subscript $i$ indicates regime 1.

In chapter 3 (equation (3.19)) we derived the utility maximising choice of leisure for the CES case and by substituting in $\sigma = 1$ we also have the Cobb-Douglas case (C-D case from now on).30 Leisure in the healthy case and in the case of illness for those individuals who take out leisure in post-recuperation time (i.e. $L \geq s r$) is then given by the following expressions:

$$L = (1 - \psi)(T_C^S) \quad \text{for } \lambda \geq 0, \lambda \neq 1$$

---

30 The maximisation yields the following first order conditions for the CES and the Cobb-Douglas case, respectively:

$$-\alpha w[w(1-L)-a]^{\lambda^2} + (1-\alpha)L^{\lambda^2} = 0, \lambda \geq 0, \lambda \neq 1$$

$$-\frac{\alpha w}{w(1-L)-a} + \frac{1-\alpha}{L} = 0, \lambda = 1$$
Define \( l - \psi = \frac{w}{\psi w^\sigma + w} \), \( \psi = \frac{\gamma w^\sigma}{\gamma w^\sigma + w} \) and \( \gamma = \left( \frac{\alpha}{1 - \alpha} \right)^\sigma \).

Notice that if the healthy individual has a minimum consumption requirement on wages that is equal to the income he earns when working full time then the individual will not spend any time in leisure. The weighting of the "choice" time in terms of leisure and labour in the C-D case are given by \( l - \alpha \) and \( \alpha \), respectively, and by \( l - \psi \) and \( \psi \) in the more general CES case.

Based on equations (4.5), (4.6) and (4.6') income in the healthy case and the regime 1 illness case is given by the following expression:

\[
y = \psi w(T_C) + a \quad \text{for } \lambda \geq 0, \lambda \neq 1 \quad (4.7)
\]

\[
y = \alpha w(T_C) + a \quad \text{for } \lambda = 1 \quad (4.7')
\]

The amount of income earned is therefore the sum of the necessary amount needed to fulfil the minimum consumption requirement, \( a \), and the utility-weighted additional possible amount to be earned.

Substituting total choice time in the healthy scenario and in the regime 1 illness scenario into equations (4.7) and (4.7'), and consequently subtracting the latter from the former yields the expression for absolute income loss (AIL). Dividing the AIL by healthy income gives us the expression for the relative income loss (RIL). Note that the subscript 1 denotes that we are dealing with regime 1, i.e. those individuals who have a wage-rate that allows them to recuperate fully:

\[
AIL_1 = y^H - y^S = \psi w r (1 - s) \quad \text{for } \lambda \geq 0, \lambda \neq 1 \quad (4.8)
\]

\[
\frac{\partial AIL}{\partial w} = r(1 - s)\psi (1 + \epsilon_{\psi, w}) = r(1 - s)\psi (\psi + \sigma (1 - \psi)) > 0 ,
\]

\[
\frac{\partial AIL}{\partial r} > 0 , \quad \frac{\partial AIL}{\partial s} < 0 , \quad \frac{\partial AIL}{\partial \alpha} = 0 , \quad \frac{\partial AIL}{\partial \sigma} > 0 , \quad \frac{\partial AIL}{\partial \sigma} = ?
\]

\[
AIL_1 = \alpha w r (1 - s) , \quad \frac{\partial AIL}{\partial w} = r(1 - s)\alpha > 0 \text{ for } \lambda = 1 \quad (4.8')
\]

\[
RIL_1 = 1 - y^S / y^H = \frac{\psi w r (1 - s)}{\psi (w - a) + a} \quad \text{for } \lambda \geq 0, \lambda \neq 1 \quad (4.9)
\]

---

See appendix 4.A1 for calculations of the differentials of equation (4.8), (4.8'), (4.9), and (4.9').
Absolute income loss increases with higher wages for those individuals who have a wage above the critical wage. In the C-D-case the reason for the higher absolute income loss of those with higher wages is easily explained. It is due to the fact that the efficient work time loss is the same for all within this group because the weight on choice labour is not affected by the wage rate, and therefore the difference in income loss is determined by the wage rate with which each hour is paid. In the more general CES-case work time loss decreases in wages if the elasticity of substitution is low (\(\sigma < 1\)) and increases in wages when it is high (\(\sigma > 1\)). Nevertheless, income loss (which is the equivalent of consumption loss) from illness always increases with higher wages for this wage group. At a low elasticity of substitution between consumption and leisure the income loss is lower and at a higher elasticity the income loss is higher.

Relative income loss, which is identical to the relative loss in efficient working hours, is higher for higher wages due to the fact that the amount of working-hours lost relative to the total amount of working hours is high for high wages, whereas the relative income loss is lower for lower wages since the hours lost is a smaller amount compared to the overall amount of hours worked. As was the case for absolute income loss, relative income loss increases in the elasticity of substitution.

As can be expected, both relative and absolute income losses increase with longer recuperation periods and decrease with higher recuperation productivity levels.

Relative and absolute income losses increase with higher exponentials (share) on consumption in the utility function. This is intuitively quite clear, since if consumption did not matter for utility there would be no income except the required minimum constraint and therefore no income loss.

The minimum consumption requirement on wage income has no effect on absolute income loss and a negative effect on relative income loss. The former follows

---

\(\frac{\partial RIL}{\partial w} = \frac{r(1-s)\alpha \psi}{(\psi(w-a)+a)^2} \frac{[\psi(1-\psi)]}{(\psi(w-a)+a)^2} > 0,\)

\(\frac{\partial RIL}{\partial r} > 0, \frac{\partial RIL}{\partial a} < 0, \frac{\partial RIL}{\partial \sigma} < 0, \frac{\partial RIL}{\partial \alpha} > 0, \frac{\partial RIL}{\partial \sigma} = 0\)

\(RIL_1 = \frac{\alpha w(1-s)}{\alpha(w-a)+a} \frac{\partial RIL}{\partial w} = \frac{r(1-s)\alpha(1-\alpha)}{(\alpha(w-a)+a)^2} > 0 \text{ for } \lambda = 1 \quad (4.9')\)

\[\frac{\partial WTL}{\partial w} = \frac{r(1-s)\psi(\sigma-1)}{\psi^{\sigma} + w} \]
directly from the fact that we are analysing the type 1, i.e. those who can afford to take more than the recuperation time out in leisure because they are not bounded by the minimum consumption requirement on wage earnings. The latter is due to the fact that the higher the minimum consumption requirement on wages, the higher is the income of the income groups we are concerned with here, and hence the loss is smaller relative to 'healthy' income.

Finally, we find that absolute and relative income losses are increasing in the elasticity of substitution if the weight on labour out of 'choice' time increases with the elasticity of substitution. This will be the case as long as the following condition holds: \( \alpha \omega > 1 - \alpha \). In words, if the relative price of leisure to consumption, \( \omega \), is larger than the relative share of leisure to consumption in total utility, \( (1 - \alpha)/\alpha \), increasing the substitutability will lead to an increase in consumption input relative to leisure.

Let us for later purposes of comparison derive analytically the lower boundaries to the income losses by substituting in the wage rate for a particular elasticity of substitution. As an example we look at the case where \( \sigma = 1 \), i.e. the Cobb Douglas utility function, and assume that \( \alpha \sigma r \leq (1 - r)(1 - \alpha) \). Substituting in the expressions for leisure time and total time into the formula in equation (3.21) from chapter 3, the critical wage is given by the following expression:

\[
\tilde{w} = \frac{a(1 - \alpha)}{(1 - r)(1 - \alpha) - \alpha \sigma r}
\]  

(4.10)

Using the expression derived for the absolute and relative income loss of the high-wage individuals (equations (4.8') and (4.9'), respectively) and substituting in the value of the elasticity of substitution and of the critical wage we obtain the lower boundaries of the absolute and relative income loss:

\[
AIL_t \geq \frac{a \sigma r (1 - s)(1 - \alpha)}{(1 - r)(1 - \alpha) - \alpha \sigma r}
\]

\[
RIL_t \geq \frac{a \sigma r (1 - s)}{1 - r + a \sigma (1 - s)}
\]

With this lower boundary to income losses derived, we will now calculate the income losses of the lower income groups in order to analyse which income groups experience the larger income losses, absolutely and relatively, due to a certain potential illness.

The only difference between the income losses derived for these higher-wage individuals in the endogenous recuperation model and the income losses of all wage-groups in the exogenous-recuperation model is that the parameter of productivity during
illness, \( s \), is not relevant in the latter and can be put equal to zero. Hence, both the absolute and relative income losses are larger by a factor of \( 1/(1-s) \), just as we found in the efficient time loss case.

### 4.3.2 Illness Case 1

For those individuals who can afford to take leisure time in the post-recuperation phase, the analysis was made in the previous section. In addition to assuming a general CES equation, let us now assume a functional form for the post-recuperation productivity:

\[
\sigma(x) = (x + (1-x)s^{1/\beta})^\beta. 
\]

This expression fulfils the conditions we assumed initially, i.e. \( \sigma(0) = s; \sigma(1) = l, s \leq \sigma(x) \leq l \) when \( 0 \leq x \leq 1 \), and gives us positive and diminishing returns to recuperation time, i.e.:

\[
\sigma'(x) = \beta(x + (1-x)s^{1/\beta})^{\beta-1}(1-s^{1/\beta}) > 0; \\
\sigma''(x) = \beta(1-s^{1/\beta})^2(\beta-1)(x + (1-x)s^{1/\beta})^{\beta-2} < 0 
\]

(here \( \sigma'(0) = \beta s^{(1-1)\beta - s^{1/\beta}}; \sigma'(1) = \beta(1-s^{1/\beta}) \)).

In case 1 derived in chapter 3, section 3.3 (\( \sigma'(l) = \beta(1-s^{1/\beta}) \geq s r/(1-r) \)) we found that the maximum amount of labour hours was achieved when exactly recuperating, i.e.:

\[
\Phi(L)_{2,1} = 1 - r 
\]

for any \( 0 \leq L \leq s r \).

Subscript 2,1 indicates that we are looking at low-wage individuals (i.e. regime 2) for illness case 1. These individuals then get the maximum amount of leisure by choosing not to work during the recuperation time and taking no leisure time in the post-recuperation phase, i.e.:

\[
L^s_{2,1} = sr 
\]

Multiplying efficient labour (equation (4.11)) by the wage rate, income takes the form:

\[
y^s_{2,1} = w(1-r) 
\]

The expressions for absolute and relative income loss can further be derived based on equations (4.7), (4.7'), and (4.13):\(^{33}\)

\[
AIL_{2,1} = wr - (1-\psi)(w-a) 
\]

for \( \lambda \geq 0, \lambda \neq 1 \)

\[
\frac{\partial AIL}{\partial \psi} = \psi + \frac{\partial \psi}{\partial w}(w-a) - (1-r) = ? \\
\frac{\partial AIL}{\partial w} = w > 0, \frac{\partial AIL}{\partial a} = 0, \frac{\partial AIL}{\partial \alpha} = (1-\psi) > 0, \frac{\partial AIL}{\partial \sigma} > 0, \frac{\partial AIL}{\partial r} = ?
\]

\(^{33}\) See appendix 4.A.2 for calculations of the differentials of equation (4.14), (4.14'), (4.15), and (4.15').
From the derivative of the absolute income loss (AIL) with respect to wages in the C-D case, we observe that AIL is decreasing (increasing) in wages if the weight given to labour out of the available "choice" time when healthy, \( \alpha \), is smaller (larger) than the amount of efficient labour supplied when ill, here; \( l - r \). In other words, if the amount of additional income made under healthy conditions with an increase in the wage rate is smaller than the additional income that would be made under illness, then the absolute income loss due to illness will decrease with increasing wages. In the CES case, the additional effect of wage on the weight given to labour must be taken into account. The absolute income loss is larger (smaller) the lower the wage only if the following (opposite) condition holds:

\[
\frac{\partial RIL}{\partial w} < 0, \quad \frac{\partial RIL}{\partial r} > 0, \quad \frac{\partial RIL}{\partial \alpha} = 0, \quad \frac{\partial RIL}{\partial \sigma} > 0, \quad \frac{\partial RIL}{\partial \sigma} = ?
\]

\[
RIL_{2,1} = \frac{\alpha(w-a) + a - w(1-r)}{\psi(w-a) + a}, \quad \frac{\partial RIL}{\partial w} = -\frac{(1-r)(1-\alpha)a}{(aw + (1-\alpha)a)^2} < 0, \quad \text{for } \lambda = l \quad (4.15')
\]

Changes in the wage rate affect the weighting of the available choice time. We therefore find that if the weight given to labour plus the relative change in this weight due to a relative change in the wage rate (i.e. the wage elasticity of the weight, \( \varepsilon_{\psi/w} \)) multiplied by the available "choice" time and the weight, is smaller (larger) than the efficient work time during illness \( (l - r) \), then AIL decreases in wages. The smaller the elasticity of substitution between consumption and leisure, the smaller will be the left hand side (LHS) of the inequality (4.16) above.

The relative income loss is larger for regime 2 case 1 the lower the wage within this group due to the fact that income after illness is compared to lower initial income in the case of the poorer, and the minimum consumption requirement on wage earnings therefore matters more.

The longer the recuperation period the higher the absolute and relative income loss, whereas the severity of the illness has no effect on either. These findings are both
due to the fact that we are looking at the type of individuals who cannot take any post-
recovery time out in leisure coupled with a type of illness where it is sub-optimal to
work during the full recovery time. The longer the recovery time lasts the more
income they lose since they will not work during this period. Having recovered when
they start working again, the severity of the illness has no effect on income.

A higher minimum consumption requirement on wage earnings increases the
income loss because it means that individuals with higher wages now belong to this
group. Relative and absolute income losses increase with higher exponentials (weight) on
consumption in the utility function. The reason for the latter finding is that with a higher
weight on consumption the individual would have worked more efficient hours in the
healthy scenario, whereas for the wage group we are considering here the amount of
hours worked during illness is constant. The weight on consumption will, however,
determine the relevant wage group, as will be clear below.

As noted earlier, locating at the kink will only be chosen when the wage-rate lies
above the existence wage, $w_{min}$, and below the critical wage, $w^*$, (based on equations
(3.23) and (3.21) in chapter 3, respectively):\textsuperscript{34}

$$\frac{a}{l-r} = w_{min} \leq w \quad \text{and} \quad w \leq w^* \quad \text{where} \quad w(l-r) - sr\gamma w^\sigma - a = 0$$

It is worth noting that if the individual in this scenario has a minimum
consumption requirement on wages that is equal to full time wages (i.e. $w = a / (l - r)$),
then the individual will still spend the entire recuperation time recuperating.

We can derive the upper and lower income loss boundaries in the case of Cobb-
Douglas preferences. With an elasticity of substitution of 1 the wage has to lie in the
following interval:

$$\frac{a}{l-r} \leq w \leq \frac{a(l-\alpha)}{(l-r)(1-\alpha)-sr\alpha} \quad (4.17)$$

Substituting in the wage rate boundaries from equation (4.17) into equations (4.14')
and (4.15'), we obtain the following analytical boundaries to absolute and relative income
losses:

\textsuperscript{34} Differentiating the expression for maximum labour when $L \geq sr$ ($\Phi(L) = sr + l - r - L$) at the point where
$L = sr$ we find that $\Phi'(sr) = -1$. However, differentiating the expression for maximum labour when $L \leq sr$
($\Phi(L) = l - r$) at the point where $L = sr$ we obtain $\Phi'(sr) = 0$ ($\geq -1$). This implies that $\Phi(L)$ is non-
differentiable at the full recuperation point $L = sr$. The individuals with wages such that the slope of their
utility curve, the marginal rate of substitution between consumption and leisure, is less than the slope of the
budget line (determined by the ratio of effective time in leisure and labour) for $L > sr$ and larger than the
slope of the budget line for $L < sr$ will locate their choice of labour and leisure exactly at the kink.
for $1 - r$ (4.18) 

$$\frac{a \alpha r}{1 - r} \geq AIL_{2,1} \geq \frac{a \alpha r(1 - s)(1 - \alpha)}{(1 - r)(1 - \alpha) - s \alpha}$$

for $\alpha < 1 - r$ (4.18) 

$$\frac{\alpha r}{1 - r + \alpha r} \geq RIL_{2,1} \geq \frac{\alpha - \alpha(1 - r)(1 - s)}{1 - r + \alpha r(1 - s)}$$

(4.19) 

Hence, if the size of the exponential of consumption in the utility function is smaller than the duration of the post-recuperation period, we will observe a decrease in absolute income loss with increasing wages until the critical wage is reached. At this point the income losses will start increasing again. If, on the other hand, the opposite inequality condition holds, we will see a continuous increase in absolute income losses with increasing wage. As for the relative income loss, it will decrease with increasing wages until the critical wage is reached and then the losses will start increasing again.

We can also compare the income losses for the low-wage individuals in the illness case 1 with the income losses for the same wage group if recuperation were modelled as exogenous. Using the absolute income loss derived in section 4.3.3, expression (4.8'), remembering that the productivity in illness can be assumed to be equal to 0 in this case, and substituting in for the relevant wage interval given in expression (4.17), we get the following AIL interval:

$$\frac{a \alpha r}{1 - r} \leq AIL_X \leq \frac{a \alpha r(1 - \alpha)}{(1 - r)(1 - \alpha) - s \alpha}$$

(4.20) 

Note that the subscript $X$ merely indicates that we are looking at the losses in the exogenous illness case. At the existence wage the individual has to work all the time possible, and given that neither illness case 1 nor the exogenous-recuperation model allows for working during the recuperation time this implies that the lower boundary to the absolute income loss is the same in both cases. However, when it comes to the upper wage-boundary we see that in the exogenous-recuperation case the AIL is higher than at the lower wage-boundary, whereas in illness case 1 the AIL may be lower or higher but it is certainly less than the AIL at the upper wage-boundary for the exogenous-recuperation case. The reason for this finding is quite clear, at the upper boundary we have those individuals who freely choose to recuperate fully in the illness case 1 and therefore in both cases efficient labour is determined by preferences for consumption in the utility function and total efficient time available. Total available time is less in the exogenous recuperation model by a factor of $(1 - s)$, and the AIL is therefore larger by a factor of $1/(1 - s)$. 

76
4.3.3 Illness Case 2

The analysis of the absolute and relative income losses for the type 1 regime was derived in section 4.3.1. For those individuals who choose not to take any post-recuperation time out in leisure, the illness in case 2 was defined by the following condition in chapter 3, section 3.3: \( \frac{sr}{1-r} \geq \sigma'(0) = \beta s^{(\beta-1)/\beta} (1-s^{1/\beta}) \). The expressions for maximum working time for given amounts of leisure were derived in chapter 3:

\[
\Phi(L) = sr - L + (1-r)\sigma\left(\frac{L}{sr}\right) \quad \text{for } 0 \leq L \leq sr \quad (4.21)
\]

\[
\Phi(L) = 1-r \quad \text{for } L = sr \quad (4.21')
\]

Let us first analyse the case where \( 0 \leq L \leq sr \). Maximising utility with respect to the budget constraint gives the following first order conditions:

\[
\frac{1-r}{sr} \sigma\left(\frac{L}{sr}\right) - 1 \left[ w (sr - L + \sigma\left(\frac{L}{sr}\right)(1-r)) - a \right]^{\lambda} + (1-\alpha) [L]^{\lambda} = 0, \text{ for } \lambda \geq 0, \lambda \neq 1 \quad (4.22)
\]

\[
L = \frac{1-\alpha}{(1-\alpha)w - \alpha} \left( w (sr + \sigma\left(\frac{L}{sr}\right)(1-r)) - a \right), \text{ for } \lambda = 1 \quad (4.22')
\]

Since leisure is equal to the amount of the recuperation time \( r \) that is spent recuperating \( x \) multiplied by the efficiency units of leisure during recuperation time \( s \) (i.e. \( L = sxr \)), we can substitute this expression for leisure into equation (4.22) above, and thereby solve the equation for \( x \) by the use of simulations. The resulting value of \( x \) can then be multiplied by the duration of the full recuperation time, and the efficiency units of leisure during recuperation time, to give the amount of leisure chosen.

Income under sickness now takes the form:

\[
y^{*}_{2,2} = w (sr(1-x) + \sigma(x)(1-r)) \quad (4.23)
\]

This expression for income is a function of the productivity in the post-recuperation time, which is a function of the amount of actual recuperation, which in turn is a function of the wage rate. We can therefore only derive exact expressions for the absolute and relative income losses for the individuals with wages within the boundaries of this wage group.

---

35 The maximisation problem is:

\[
\begin{align*}
\max_L \quad & \frac{\alpha}{1-\lambda} \left[ w (sr - L + \sigma\left(\frac{L}{sr}\right)(1-r)) - a \right]^{\lambda-1} + \frac{1-\alpha}{1-\lambda} [L]^{\lambda}, \text{ for } \lambda \geq 0, \lambda \neq 1 \\
\max_L \quad & \left( w (sr - L + \sigma\left(\frac{L}{sr}\right)(1-r)) - a \right)^{\lambda}, \text{ for } \lambda = 1
\end{align*}
\]
numerically. This is done in example 2 in section 4.5.2.

The expressions for absolute and relative income loss for the low-wage individuals in case 2, regime 2, (indicated by the subscript 2,2 below) can be derived based on equations (4.7), (4.7') and (4.23), in the special cases of $x=0$ and $x=1$:

$$AIL_{2,2} = \begin{cases} 
\psi(w-a) + a - ws, x = 0 \\
\psi(w-a) + a - w(1-r), x = 1
\end{cases}$$

for $\lambda \geq 0, \lambda \neq 1$  \hspace{1cm} (4.24)

$$AIL_{2,2} = \begin{cases} 
\alpha(w-a) + a - ws, x = 0 \\
\alpha(w-a) + a - w(1-r), x = 1
\end{cases}$$

for $\lambda = 1$  \hspace{1cm} (4.24')

$$RIL_{2,2} = \begin{cases} 
\psi(w-a) + a - ws, x = 0 \\
\psi(w-a) + a - w(1-r), x = 1
\end{cases}$$

for $\lambda \geq 0, \lambda \neq 1$  \hspace{1cm} (4.25)

$$RIL_{2,2} = \begin{cases} 
\alpha(w-a) + a - ws, x = 0 \\
\alpha(w-a) + a - w(1-r), x = 1
\end{cases}$$

for $\lambda = 1$  \hspace{1cm} (4.25')

By substituting for $x = 0$ in equation (4.22) and (4.22') and finding the wage that makes the two sides of the equation equal we find the lower bound to the wages that make this solution optimal:

$$w = \frac{a}{s}$$

(4.26)

Similarly, by substituting for $x=1$ in equation (4.22) and finding the wage that equalises the two sides we find the upper bound to the wages, $\tilde{w}$ (also referred to as the lower kink-wage):

$$\frac{\alpha \tilde{w}}{1-\alpha} \left(1 - \frac{1-r}{sr} \right) = \left[ \frac{\tilde{w}(1-r)-a}{sr} \right]^\lambda$$

(4.27)

It is important to note that $\Phi(L)$ is non-differentiable at the full recuperation point $L = sr$.

The individuals with wages such that the slope of their utility curve, the marginal rate of substitution between consumption and leisure, is less than the slope of the budget line for $L > sr$ and larger than the slope of the budget line for $L < sr$ will locate their choice of labour and leisure exactly at the kink.

A particular case where we can derive the upper wage boundary is the one of Cobb-Douglas preferences (i.e. $\sigma = \lambda = 1$). Substituting for $\lambda = 1$ in equation (4.27) the range within which the wage rate lies is the following:
Substituting the wage boundaries into the expressions (4.24') and (4.25'), the following boundaries for absolute and relative income losses result:

\[
\text{AIL}_{2,2} = \begin{cases} 
\frac{\alpha}{1 - \lambda}, x = 0 \\
\frac{\alpha}{\lambda (1-s) + \lambda'(1)(1-r)}, x = l
\end{cases}
\text{for } \lambda = 1
\]

\[
\text{RIL}_{2,2} = \begin{cases} 
\frac{1 - s}{\alpha + s(1-\alpha)}, x = 0 \\
\frac{1 - r}{\lambda(1-sr) + (1-r)(1-\alpha + \lambda'(1))}, x = l
\end{cases}
\text{for } \lambda = 1
\]

Comparing the expressions for AIL when wage is low (here: \(x = 0\)) and wage is higher (\(x = l\)) given in expression (4.29), we find that the AIL is lower at the higher wage than the lower wage if the following condition holds:

\[
\frac{s}{1-r} < \frac{(1-s)(1-\alpha + \lambda'(1))}{\lambda(1-s) + \lambda'(1)(1-r)(1-\alpha)}
\]

The relative income loss (RIL) is lower at the higher wage if the following condition is fulfilled:

\[
\frac{s}{1-r} < \frac{1}{1-sr + (1-r)\lambda'(1)}
\]

Keeping in mind that for illness case 2 the productivity during illness, \(s\), is larger than the duration of the post-recuperation period, \(1-r\), these conditions imply that the numerators on the right hand side (RHS) are larger than the denominators. Hence, we see that whether AIL and RIL decrease or increase when moving from the lower to the higher wage depends on the characteristics of the illness, i.e. the curvature of the marginal productivity in the post-recuperation period of additional recuperation, \(\beta\), its duration, \(r\), its severity, \(1-s\), as well as on the taste of the individual (i.e. the exponentials in the utility function, \(\alpha\)).

Locating at the kink, where \(L - sr\), will only be chosen when wage lies above the wage, \(\tilde{w}\), and below the critical wage, \(\bar{w}\), (based on equations (4.27) and (4.18)):

\[
w \geq \tilde{w} \quad \text{where} \quad \frac{\alpha \tilde{w}}{1-\alpha} \left(1 - \frac{1-r}{sr}\lambda'(l)\right) = \left[\frac{\tilde{w}(1-r) - \alpha}{sr}\right]^{\frac{1}{\lambda}} \quad \text{and}
\]

\[
w \leq \bar{w} \quad \text{where} \quad \bar{w}(1-r) - sr\tilde{w}^{\lambda} - \alpha = 0
\]
For example, if the elasticity of substitution is 1 (i.e. Cobb-Douglas preferences) the wage has to lie in the following interval:

\[ \frac{a(1-\alpha)}{(1-r)(l-\alpha + a\sigma'(1\rangle) - s\alpha} \leq w \leq \frac{a(1-\alpha)}{(1-r)(1-\alpha) - s\alpha} \quad (4.31) \]

The expressions for income and absolute and relative income losses for this leisure group were derived under illness case 1 (equations (4.13), (4.14), and (4.15), respectively). The relative income loss was found to be larger the lower the wage within this group, whereas the absolute income loss was found to be larger the lower the wage only if the weight given to consumption in total utility (which in the general CES-case is influenced by the wage-rate) is smaller than the efficient work time during illness, \( l - r \). See section 4.3.2 for more details.

Substituting in the expression for the wage rate from equation (4.31) into equations (4.14') and (4.15') we obtain the following analytical boundaries for absolute and relative income losses:

\[ \frac{a(1-\alpha)[r(1-s)+(1-r)\sigma'(1\rangle)}{(1-r)(1-\alpha + a\sigma'(1\rangle) - s\alpha} \geq AIL \geq \frac{a(1-\alpha)r(1-s)}{(1-r)(1-\alpha) - s\alpha} \quad \text{for } \alpha < 1-r \quad (4.32) \]

\[ \frac{\alpha - \alpha(1-r)(1-\alpha) - asr}{\alpha + (1-r)(1-\alpha + a\sigma'(1\rangle) - asr} \geq RIL \geq \frac{\alpha - \alpha(1-r)(1-s)}{1-r + \alpha r(1-s)} \quad (4.33) \]

We can also compare the income losses for the low-wage individuals in the illness case 2 with the income losses for the same wage group if recuperation were modelled as exogenous. Using the absolute income loss derived in section 4.3.1, expression (8'), remembering that the productivity in illness can be assumed to be equal to 0 in this case, and substituting in for the relevant wage interval given in expression (4.31), we get the following AIL interval:

\[ \frac{a\sigma(1-\alpha)}{(1-r)(1-\alpha + a\sigma'(1\rangle) - s\alpha} \leq AIL \leq \frac{a\sigma(1-\alpha)}{(1-r)(1-\alpha) - s\alpha} \quad (4.34) \]

Both at the lower kink-wage and at the critical wage the AIL is higher in the exogenous recuperation case, however the difference in income loss predicted by the two models is larger for those individuals receiving the critical wage than for those receiving the lower kink-wage. Comparing equations (4.32) and (4.34) the AIL is larger by a factor of \( (1-s)^{-1} \) at the upper boundary and by a factor of \( \left[ \frac{(1-s)+\frac{(1-r)\sigma'(1\rangle}{r} \right]^{-1} \) at the lower. The reason for this finding is that at the upper boundary we have those individuals who freely choose to recuperate and therefore in both cases efficient labour is determined by preferences for consumption in the utility function and total efficient time available. At
the lower boundary the choice of labour is still free in the case of the exogenous-recuperation model, whereas in the endogenous recuperation model the individuals would have wanted to trade some more leisure for labour had the terms of trade still been one to one, but they are not and hence their choice is more restricted. When comparing income losses for those individuals with a wage-rate below the lower kink-wage, it is worth noting that the existence wage is lower in the case of the endogenous recuperation model than in the exogenous one - in other words, poorer people can survive in the former case. Comparing the AIL for the individuals with a wage between the exogenous-recuperation model existence wage and the lower kink-wage, it is clear that the difference in income loss decreases with lower wages. However, unlike in illness case 1, even at the lowest possible wage-rate the AIL is here still larger for the exogenous-recuperation model.

4.3.4 Interim Summary

In measuring the income loss due to illness, we have once again seen that the traditional way of modelling illness as time lost will seriously misrepresent the adverse effect of illness if the duration of illness is in fact a choice variable. The exogenous recuperation model predicts that both absolute and relative income losses increase with increasing wages.

In the case of a relatively severe illness (case 1) where there is nothing to be gained by working while being sick, however, relative income loss in the endogenous recuperation model is found to be decreasing with increasing wages until the critical wage is reached (where the marginal rate of substitution between consumption and leisure is equal to the "healthy" part of the efficient time budget constraint), and then the losses start increasing again. Furthermore, we find in the case of a C-D utility function, that if the size of the exponential of consumption is smaller than the duration of the post-recuperation period, we will observe a decrease also in absolute income losses with increasing wages until the critical wage is reached and then the losses will start increasing again. In figure 4.6 this case has been illustrated and compared to the income losses predicted by the exogenous recuperation model (EXOR). If, on the other hand, the opposite inequality condition holds, we will see a continuous increase in absolute income losses with increasing wage. Hence, there is a possibility that the individuals with the lowest wages loose more income, both absolutely and relatively, than individuals with higher wages (up to a certain wage level) due to the removal of the opportunity to earn additional income during illness.
In the case of a not so severe illness (illness case 2), where low-wage individuals choose to work part of their recuperation time, both absolute and relative income losses may be either increasing or decreasing with increasing wages up to the lower boundary kink-wage, depending on the characteristics of the illness and the taste of the individual. For those wage-groups that choose to exactly fully recuperate, relative income losses will decrease with higher wages and so will absolute income losses if the size of the exponential of consumption in the utility function is smaller than the duration of the post-recuperation period (same as for illness case 1). With wages above the critical wage both the absolute and relative income losses will start increasing again. Hence, there is also in this illness case a possibility that the individuals with the lowest wages loose more income, both absolutely and relatively, than individuals with higher wages (up to a certain wage level) due to the constraint on the opportunity to earn additional income during illness. This case has been illustrated in figure 4.7 and contrasted to the prediction of the EXOR-model.

In the illness case 1 the loss of opportunity to earn additional income during the illness period is total, whereas in the second illness case it is only partial; individuals can choose to work during illness but at a lower productivity and at the cost of an extended illness period. Therefore the adverse health effect is worse in the first illness case, and this is supported by the income loss measurement.

The minimum wage-rate required in order to survive, the existence wage, is lower in the illness case 2 than 1, implying that poorer people are able to survive when the illness is less serious because they are able to earn additional income during their illness. For any wage rate below the lower boundary “kink-wage” in the second illness case and above the existence wage in the first illness case, the absolute income loss will be larger in illness case 1 than in illness case 2.

We have seen that under certain conditions the lowest-wage individuals may loose more income (even absolutely) than higher wage individuals due to the same illness (identical illness characteristics), and illnesses may thereby have income-distributional implications. However, if we want to have a proper measurement of the full adverse effects of illness we will also have to take into consideration that consumption is only a part of welfare. Another factor usually assumed to provide welfare to the individual is leisure time, and a true measurement of the adverse health effect must therefore take both these factors into account. In the next section we will derive the welfare loss due to illness of different wage-groups.
Figure 4. 6: Income losses (absolute and relative) at various wage rates for illness case 1 and the EXOR model.

Figure 4. 7: Income losses (absolute and relative) at various wage rates for illness case 2 and the EXOR model.

Note: The figures above are based on the numerical examples given in section 4.5 of this chapter. We have assumed a duration of the full recuperation, $r$, of 0.3 (out of full time which is unity), a weight on consumption in the utility function, $\alpha$, of 0.5, and a net minimum consumption requirement, $\theta$, of 1. In addition we have assumed the productivity during illness, $s$, in the case 1 to be 0.5 and in the case 2 to be 0.8.
4.4 Welfare Loss from Illness

So far in this chapter we have seen that the traditional statistical measurement of illness of work-absence is very misleading because it does not distinguish between chosen and imposed work-absence and, more importantly, does not capture the illness of those who work while being ill. The illness-experience of the latter is better captured by the measurement of efficient time loss, as this looks at productivity losses no matter what activities the individual is undertaking. However, efficient time loss does not perform well in the case of illnesses where productivity is so low that no additional consumption is gained by working during illness. With this measurement of adverse health effects the low-wage individuals will appear to have lost as much as those individuals who choose not to work during recuperation, whereas their loss is in some sense much more serious since the opportunity to earn income during illness has been removed. The removal of the opportunity to work is better picked up by our calculation of income loss. However, this latter measurement also suffers from one disadvantage: it does not capture the direct effect of illness on the wellbeing of an individual. This is why we now turn to the calculation of welfare loss.

As mentioned in chapter 2, Preston and Walker (1992) review a number of welfare measures that can be used in the context of non-linear budget constraints. The full income expenditure function at a reference wage has however been chosen here since it is the classic money metric of utility, which can be computed and interpreted easily, and directly compared with net income.

The following analysis draws on King (1983) and Aaberge and Colombino (1998) by deriving measures of welfare from equivalent incomes defined in terms of a reference household (here; individual) and the prices (here; reference wage) this household faces. The introduction of a common reference wage as a basis for comparing welfare across individuals is motivated by the fact that real/efficient wage, the price of leisure, will vary across individuals as well as between states (here; healthy state and state of illness).

Although we have distinguished between wage and efficient time in the preceding chapters, it is important to realise that efficient wage is the wage received per nominal hour (in contrast to efficient hour) and may therefore vary between the healthy and sick state, as well as between non-poor and poor individuals.

---

36 See chapter 2, section 2.2.2, for a review of possible welfare measures that can be used in the context of
The indirect utility function of full income and wage is given by maximising the direct utility, which is a function of consumption and leisure, subject to the budget constraint:

\[ v(w, F) = \max_u u(C, L) \text{ s.t. } C + wL \leq F \]

where \( v \) is the indirect utility function, \( w \) is wage, \( F \) is full income, \( u \) is the direct utility function, \( C \) consumption, and \( L \) leisure.

The approach for determining the change in welfare of one particular individual due to illness is to employ monetary measures defined in terms of money values of indirect utilities. For given consumer prices and a wage the money metric utility, \( F_i^k \), is defined implicitly by

\[ V_i^k(F_i^k, w_i) = \max_L u[w_i\Phi^k(L) - a_iL], \quad k = h, s \quad \text{and} \quad i = r, p \]

where \( V_i \) is the indirect utility function of individual \( i \) and \( F_i^k \) is the full income of individual \( i \) in the state of health \( k \). The wage and the minimum consumption requirement on wage earnings will vary over individuals, \( i \). We will in addition distinguish between two states of health, \( k \). The state of health denoted \( h \) implies that the individual is healthy the entire period, and the one denoted \( s \) that part of the period is spent being sick. Thus, the full income \( F_i^k \) affords individual \( i \) the same level of indirect utility under the wage \( w_i \) as the maximum amount of obtainable direct utility attained with wage \( w_i \) with full income \( F_i^k \).

In order to make these money measures of utility comparable across individuals King suggested to base the comparison on equivalent incomes defined in terms of a reference household (here; individual) facing a reference price. Since the only factor separating different individuals in this model is the wage rate, this method implies in practice using the same reference wage for all individuals. Equivalent income, \( F_i^k \), for individual \( i \) is then defined as that level of full income that yields the same level of utility at the reference wage, \( \bar{w} \), as the same individual \( i \) attains under the wage \( w_i \). \( F_i^k \) is given implicitly by the following expression:

\[ v(\bar{w}, F_i^k) = V_i^k = \max_L u[w_i\Phi^k(L) - a_iL], \quad k = h, s \quad \text{and} \quad i = r, p \]
From this expression we see that the difference between $F_i^h$ and $F_i^s$ can be used as a measure of the welfare loss from illness to individual $i$, and as the money values are defined in terms of a fixed reference wage, this choice of measure allows for welfare change comparisons across individuals. We denote the measure the *absolute comparable welfare loss* (ACWL):

$$ACWL_i = F_i^h - F_i^s$$  \hspace{1cm} (4.35)

The *relative comparable welfare loss* (RCWL) will therefore be given by the following expression (relative to the healthy state):

$$RCWL_i = 1 - \frac{F_i^s}{F_i^h}$$  \hspace{1cm} (4.36)

Both King (1983) and Aaberge et al. (1998) mention the possibility that the outcome of the comparison of welfare changes may depend on the choice of reference price, and it will be important to examine the sensitivity of the results with regards to the choice of reference wage. It is however worth noting that Aaberge et al. carry out such a sensitivity analysis and find that the main conclusions are not affected by the choice of reference state.

In order to examine some results more explicitly we have once more to return to our CES utility function:

$$U(c,L) = \begin{cases} 
\frac{\alpha}{1 - \lambda} [c]^{\lambda-1} + \frac{1 - \alpha}{1 - \lambda} [L]^{\lambda-1}, & \lambda \geq 0, \lambda \neq 1 \\
\alpha \log(c) + (1 - \alpha) \log(L), & \lambda = 1
\end{cases}$$

where $C$ is consumption and $L$ is leisure, and where $0 < \alpha < 1$, and $\sigma = \frac{1}{\lambda}$ is the elasticity of substitution. The Cobb-Douglas case is where $\sigma = \lambda = 1$.

The budget constraint is given by the following expression:

$$C + L w \leq T(L) w - a \quad \text{and} \quad F = T w - a$$  \hspace{1cm} (4.37)

where full income, $F$, is equal to total time, $T$, multiplied by the wage rate, $w$, minus $a$, the minimum consumption requirement on wage-earnings. The price of consumption, $p^C$, has been normalised to one.

In the analysis below the absolute and relative welfare losses due to sickness for the different wage groups will be derived, both for the general CES case and the C-D case. However, when explaining the marginal effects of the different parameters and
variables on the size of the losses we will concentrate on the C-D case, as the differentials obtained using this preference structure are more easily interpreted and explained.

4.4.1 Welfare Loss in the Full Recuperation and the Exogenous Recuperation Cases

Indirect utility is found by substituting the utility maximising choice of leisure into the budget constraint (equation (4.37)) solving for consumption and substituting the two Marshallian demand functions for consumption and leisure back into the utility function:

\[ V^k = v(w, T^k, w - \alpha) = \left( \frac{F^k}{w} + w \right)^{1-\lambda} \left( \frac{\alpha}{1-\lambda} (\gamma w^\omega)^{1-\lambda} + \frac{1-\alpha}{1-\lambda} \right) \text{ for } \lambda \geq 0, \lambda \neq 1 \] (4.38)

with \( \frac{\partial V}{\partial w} > 0 \)

\[ V^k = v(w, T^k, w - \alpha) = (F^k)w^{-(1-\lambda)} \alpha^\omega (1-\alpha) \lambda^{-\alpha}, \quad \frac{\partial V}{\partial w} > 0 \text{ for } \lambda = 1 \] (4.38')

Full income, \( F^k \), was given by the following expression in the healthy case: \( F^H = w - \alpha \). In the sickness scenario regime 1 we had the following expression for full income: \( F^H = w(1-r(1-s)) - \alpha \). Solving equation (4.38) for full income, we can then express the income equivalent of utility at a reference wage, \( \overline{w} \), in the following manner:

\[ F^k_w = \left[ \frac{v^k}{\alpha^{\omega} (\gamma \overline{w}^{\omega} + \overline{w})} \right]^{1-\lambda} \left( \frac{\gamma \overline{w}^{\omega} + \overline{w}}{1-\lambda} \right) \] (4.39)

where \( F^k_w \) indicates the minimum amount of money in excess of the net minimum consumption requirement required at a reference wage (\( \overline{w} \)) to support the level of indirect utility \( v^k \). In the Cobb-Douglas case the income equivalent at a reference wage can be derived from expression (4.38'):

\[ F^k_w = \frac{v(w)\overline{w}^{1-\alpha}}{\alpha^\omega (1-\alpha) \lambda^{-\alpha}} \] (4.39')

By substituting the expressions for full income in the healthy and sick case into the indirect utility expressions given in equations (4.38) and (4.38'), and substituting these again into the expressions for full income equivalents of welfare at a reference wage

\[ ^{37} \text{In the cases of the healthy scenario and the regime 1 sickness scenario (i.e. those who choose to recuperate fully), the utility maximising choices of leisure were given in expression (4.6) and (4.6'), for CES and C-D preferences, respectively.} \]
(equations (4.39) and (4.39')), and then finally substituting the latter into the general expressions for absolute and relative welfare losses (equations (4.35) and (4.36)) we obtain the following expressions for absolute comparable welfare loss (ACWL) and relative comparable welfare loss (RCWL), respectively:38

\[ ACWL_j = \frac{\alpha(n^\sigma \gamma^{-\lambda} + 1 - \alpha)}{\alpha(n^\sigma \gamma^{-\lambda} + 1 - \alpha)} \left( \frac{\bar{n}^\sigma + \bar{w}}{n^\sigma + w} \right) wr(1 - s) \]  

(4.40)

where \( \frac{\partial ACWL}{\partial w} > 0, \frac{\partial ACWL}{\partial s} < 0, \frac{\partial ACWL}{\partial r} > 0, \frac{\partial ACWL}{\partial \alpha} > 0, \frac{\partial ACWL}{\partial \sigma} = 0, \frac{\partial ACWL}{\partial \alpha} = ? \), \( \frac{\partial ACWL}{\partial \sigma} = ? \)

In the Cobb-Douglas case:

\[ ACWL_j = \left( \frac{w}{w} \right)^{1-a} wr(1 - s) \]  

(4.40')

CES and Cobb-Douglas case:

\[ RCWL_j = \frac{wr(I - s)}{w - a} \]  

(4.41)

where \( \frac{\partial RCWL}{\partial w} < 0, \frac{\partial RCWL}{\partial s} < 0, \frac{\partial RCWL}{\partial r} > 0, \frac{\partial RCWL}{\partial \alpha} > 0, \frac{\partial RCWL}{\partial \sigma} = 0 \)

A higher wage leads to a higher absolute welfare loss. This is intuitively quite clear because those individuals with higher wages loose a higher amount of consumption (although the same amount of efficient labour hours), whereas the amount of leisure time lost is the same and this leads to a higher absolute welfare loss.

Relative welfare loss, on the other hand, is decreasing in wage due to the fact that welfare after illness is compared to lower initial welfare in the case of the poorer, and the minimum consumption requirement on wage earnings therefore matters more.

The length of the recuperation period has a positive effect on absolute and relative welfare losses, meaning that longer recuperation leads to larger welfare losses. Conversely, productivity during illness has a negative effect on both absolute and relative welfare losses. A higher productivity during illness leads to a lower welfare loss.

Relative welfare loss is increasing in the net minimum consumption requirement, \( a \), but the latter has no effect on absolute welfare loss. The former is due to the fact that the higher the net minimum consumption requirement, the lower is the initial welfare of the income groups we are concerned with here, and hence the loss is larger relative to

---

38 See appendix 4.A3 for calculations of the partial differentials of expressions (4.40), (4.40'), and (4.41).
'healthy' welfare. The latter follows directly from the fact that we are analysing regime I, i.e. those who take more than the recuperation time out in leisure because they are not restricted by the minimum consumption requirement on wage earnings.

It is worth noticing that neither the choice of reference wage nor the relative share of labour, $\alpha$, matter for relative welfare loss. The former is due to the reference wage constituting a unit of measurement, and clearly does not matter for the calculation of relative losses, and the latter is a weighting factor which does not change between the healthy and the sick state and therefore does not influence relative loss measurements. However, absolute welfare loss is increasing in the reference wage, meaning that a higher reference wage leads to higher losses absolutely (again due to being a unit of measurement), whereas the sign of the exponential is positive if the individual's actual wage is larger than the reference wage and is negative in the opposite case.

Substituting the critical wage (equation (4.10)) into the expressions for absolute and relative welfare loss in the case of Cobb-Douglas preferences we obtain the following boundaries for these losses:

$$ACWL_1 > \tilde{w}^{1-\alpha} r (1-s) \left( \frac{\alpha(1-\alpha)}{(1-r)(1-\alpha)-\alpha r} \right)^{\alpha}$$

$$RCWL_1 < \frac{(1-s)(1-\alpha)}{1-\alpha + \alpha s}$$

The only difference between the welfare losses derived for these higher-wage individuals in the endogenous recuperation model and the income losses of all wage-groups in the exogenous-recuperation model is that the parameter of productivity during illness, $s$, is not relevant and can be put equal to zero. Hence, both the absolute and relative income losses are larger by a factor of $1/(1-s)$, just as we found in the efficient time loss and the income loss calculations.

4.4.2 Illness Case 1

It was previously shown that individuals with wages below the critical wage, $\tilde{w}$, choose not to fully recuperate. In the illness case 1, however, there was no additional efficient labour to be gained by working during the recuperation time. Hence they would locate their choice of leisure time exactly at the kink, where $L = s r$ and $H = 1 - r$. Their indirect utility is therefore:
\[
v_{2,i} = \frac{\alpha}{1-\lambda} (w(l-r) - a)^{1-a} + \frac{1-\alpha}{1-\lambda} (sr)^{1-a} \quad \frac{\partial v}{\partial w} > 0 \quad (4.44)
\]

C-D case:
\[
v_{2,i} = (w(l-r) - a)^{a} (sr)^{1-a} \quad \frac{\partial v}{\partial w} > 0 \quad (4.44')
\]

As was the case for the indirect utility expression of regime 1 individuals, \(v(w)\) is a strictly increasing function of \(w\). Furthermore, \(v(w)\) can be shown to be a continuous function of \(w\). Plugging the critical wage rate for the Cobb-Douglas case (given in equation (10)) into the expressions for indirect utility in the regime 2 (low-wage) case (equation (44')) and in the regime 1 case (equation (38')) we see that the critical wage rate satisfies the following condition:
\[
\left( \tilde{w}(1-r) - a \right)^{a} (sr)^{1-a} = \left( \tilde{w}(1-r(1-s)) - a \right)^{a} (1-\alpha)^{1-a} \tilde{w}^{-(1-a)}
\]

Thus for all wages larger or equal to the existence wage, \(w_{\text{min}} \geq a/(1-r)\), indirect utility is a strictly increasing and continuous function of \(w\). Moreover, as \(w \to a/(1-r)\), \(v(w) \to 0\). The latter finding merely reflects the fact that when the wage approaches the existence wage individual choice disappears.

The full-income equivalent measure of welfare using a reference wage rate was given in equation (4.39) for the general CES utility function and in expression (4.39') for the Cobb-Douglas case. The absolute and relative comparable welfare losses can therefore now be expressed as follows:
\[
ACWL = \frac{\tilde{w}^{a} + \tilde{w}}{\left[ \alpha (\tilde{w}^{a})^{1-\frac{1}{\lambda}} + (1-\alpha)^{1-\frac{1}{\lambda}} \right]^{1-\frac{1}{\lambda}}} \left[ \tilde{w}^{\frac{\alpha}{w}(\tilde{w}^{a})^{1-\frac{1}{\lambda}} + (1-\alpha)^{1-\frac{1}{\lambda}}} - \left( \frac{\alpha}{w}(w(1-r) - a)^{1-a} + (1-\alpha)(sr)^{1-a} \right)^{1-\frac{1}{\lambda}} \right]^{1-\frac{1}{\lambda}}
\]
\[
(4.45)
\]

C-D case:
\[
ACWL = \tilde{w}^{1-a} w^{a} \left\{ 1 - \frac{\alpha}{w} \left( \frac{1-r-a}{w} (sr)^{1-a} \right) \right\} \quad (4.45')
\]
\[
\frac{\partial ACWL}{\partial w}, \frac{\partial ACWL}{\partial r}, \frac{\partial ACWL}{\partial s}, \frac{\partial ACWL}{\partial a}, \frac{\partial ACWL}{\partial \alpha} = ?
\]

---

39 See appendix 4.A4 for the calculation of the partial derivates.
Absolute welfare losses decrease with increasing wages when the additional wage leads to more additional welfare in the sick case than in the healthy case. With C-D preferences this is the case if the optimal amount of consumption in the healthy scenario is smaller than the amount of consumption chosen under illness case 1, multiplied by the marginal rate of substitution between efficient leisure and efficient labour in the sickness scenario to the power of $(1-\alpha)$, i.e. the weight given to leisure in the utility function, i.e.:

$$\alpha \left( \frac{1-r}{w} \right) + \frac{a}{w} < \left( \frac{\alpha w s r}{(1-\alpha)(w(1-r)-a)} \right)^{1-a} (1-r)$$

If we think of efficient labour here in terms of leisure equivalents, the expression above means that when the marginal rate of substitution is infinity in the sickness case (i.e. the individual would give up any amount of efficient leisure to obtain an additional unit of efficient labour) the increase in consumption due to a marginal increase in wage would be worth indefinitely much in terms of leisure equivalents, and certainly more than an increase in consumption due to a marginal wage increase in the healthy case (which can be traded one for one with leisure). On the other hand, when the marginal rate of substitution is 1, efficient labour both in the sickness case and the healthy case can be substituted one for one with efficient labour, and in this case an increase in wage will obviously have a larger impact in the healthy scenario since the individual provides more efficient labour then. The condition above is thus fulfilled when the slope of the budget constraint is 0, but not when it is equal to -1 (at the critical wage). Some marginal rate of substitution between 0 and 1 will make the two sides of the inequality equal, i.e. at that rate of substitution the wage rate will not affect ACWL.

Relative welfare loss is decreasing with increasing wages when consumption in the healthy scenario exceeds consumption in the sickness scenario, i.e.:

$$\alpha (w-a) + a > w(1-r).$$

This condition is necessarily fulfilled since consumption is
assumed to be a normal good. Relative welfare loss is decreasing in wage due to the fact that reducing the net consumption constraint by increasing wages has a larger relative effect (positive) on welfare in the illness scenario than in the healthy scenario.

The reference wage has no effect on relative welfare loss and a positive effect on absolute welfare loss (i.e. losses are increasing in the reference wage) for the same reasons as were given under the regime one discussion. Absolute and relative welfare losses are once more increasing in the duration of the illness parameter and decreasing in the productivity during illness parameter.

Relative welfare loss is increasing in the net minimum consumption requirement, \(a\), due to the fact that the higher the net minimum consumption requirement, the lower is the initial welfare of the income groups we are concerned with here, and hence the loss is larger relative to 'healthy' welfare. However, contrary to what we found for the regime 1 individuals, the regime two individuals in illness case 1 are restricted by the minimum consumption requirement on wage earnings, and the stronger the restriction (i.e. the higher \(a\)) the higher their absolute welfare loss.

Finally, by substituting in the relevant wage boundaries we see that at a wage rate below the critical wage relative welfare loss is increasing in the weight on consumption in the utility function, as opposed to the regime 1 (high-wage) case where it had no effect. The reason why it now matters is that the opportunity to work additional efficient hours during the illness have been removed due to the nature of the illness, and the higher the weight on consumption the more the individual would have chosen to work had the opportunity been there. For this same reason absolute welfare loss will also increase in the weight on consumption, unless the reference wage is so much higher than the actual wage that their ratio, which is weighting the contribution of leisure in the full income equivalent of utility, counteracts this effect. In particular, when \(\alpha\) is increasing the weight on the leisure contribution decreases and this will therefore decrease the measure of the absolute comparable welfare loss.

When substituting in the relevant wage-boundaries in the C-D-case, we find that relative welfare loss will lie in the following interval:

\[
I > RCWL \geq I - \frac{s}{1 - \alpha(1 - s)}
\]  

(4.47)

Absolute welfare loss, on the other hand, will be higher at the existence wage than at the critical wage if the following condition holds:
The LHS of equation (4.48) gives ACWL at the existence wage and the RHS at the critical wage. We know, however, that the ACWL will decrease in wages up to a certain wage between the existence wage and the critical wage and then start increasing again.

Finally we want to compare the welfare losses for the low-wage individuals in the illness case 1 with those predicted by the exogenous illness model (ACWL<sub>x</sub>). The latter can be found by taking the expressions for absolute welfare loss in the regime 1 case, given in formula (4.40'), and substituting in for the productivity during illness, s (s=0 in the exogenous recuperation model) and the relevant wage interval given in expression (4.17). We obtain the following ACWL interval:

\[ \bar{w}^{1-a} \left( \frac{a}{1-r} \right)^{a} > \bar{w}^{1-a} r (1-s) \left( \frac{a(1-\alpha)}{(1-r)(1-\alpha)-sr} \right)^{a} \]

(4.48)

when \( \left( \frac{w_{\text{min}}}{\bar{w}} \right)^{a} = \left( 1-s \left( \frac{r}{1-r} \right) \frac{a}{1-\alpha} \right)^{a} > 1-s \)

At the existence wage the welfare loss is identical for both cases. This reflects the fact that when the wage approaches the existence wage individual choice disappears. Hence, although in the endogenous model the individuals obtain a certain amount of leisure, the fact that utility is a product of consumption and leisure means that since there is no choice consumption this leisure gives no indirect utility. However, when it comes to the upper wage-boundary, we see that in the exogenous-recuperation case the ACWL is higher than at the lower wage-boundary, whereas in illness case 1 the ACWL may be lower or higher but it is certainly less than the AIL at the upper wage-boundary for the exogenous-recuperation case. This is because in both cases efficient choice labour and leisure are now determined by preferences for consumption and leisure in the utility function and total efficient time available. Total available time is less in the exogenous recuperation model by a factor of \( (1-s)^{-1} \), and the ACWL is therefore larger by a factor of \( 1/(1-s) \).

4.4.3 Illness Case 2

In contrast to illness case 1, individuals with wages below the critical wage, \( \tilde{w} \), in the illness case 2 can obtain additional efficient labour by working during the recuperation time. In contrast to those individuals with a wage rate above the critical
wage who face a slope of the budget constraint of -1, however, the slope faced by those with low wages will be -1 or less implying that they will have to give up more leisure to obtain an additional unit of labour.

For the individuals who choose a bundle located on the curved part of the budget constraint we know that at the points of tangency between the efficient time budget constraint and the indifference curve the slope of the budget constraint is equal to the marginal rate of substitution between efficient labour and efficient leisure (negative of the ratio of the marginal utility of leisure to the marginal utility of labour), i.e.:

\[
\Phi'(L) = \frac{\partial U / \partial L}{\partial U / \partial H} = -\frac{1-\alpha}{w\alpha} \left(\frac{C}{L}\right)^{1-a} \quad \text{for } \lambda \geq 0, \lambda \neq 1
\]  

\[
\Phi'(L) = -\frac{\partial U / \partial L}{\partial U / \partial H} = -\frac{1-\alpha}{w\alpha} \frac{C}{L} \quad \text{for } \lambda = 1
\]  

(4.49)

The equation for the slope was developed in chapter 3:

\[
\Phi(L) = \frac{1}{sr} \sigma'(x) - 1 < 0
\]  

(4.50)

By solving equations (4.49) and (4.49') for leisure, substituting the expression obtained into the budget constraint, and solving for consumption we obtain:

\[
C = \frac{\gamma w^r (-\Phi')^r(wT^s - a)}{\gamma w^r (-\Phi')^r + w} \quad \text{for } \lambda \geq 0, \lambda \neq 1
\]  

(4.51)

\[
C = \frac{a(-\Phi')^r wT^s - a}{1-a - a\Phi'} \quad \text{for } \lambda = 1
\]  

(4.51')

Furthermore, substituting this expression for consumption into the expression for leisure we obtain:

\[
L = \frac{wT^s - a}{\gamma w^r (-\Phi')^r + w} \quad \text{for } \lambda \geq 0, \lambda \neq 1
\]  

(4.52)

\[
L = \frac{(1-a)(wT^s - a)}{w(1-a - a\Phi')} \quad \text{for } \lambda = 1
\]  

(4.52')

We can now calculate the indirect utility, \(v(w)\), by substituting these two Marshallian demand functions for consumption and leisure back into the utility function:

\[
v(w) = a \left(\frac{\gamma w^r (-\Phi')^r(wT^s - a)}{\gamma w^r (-\Phi')^r + w}\right)^{1-a} + (1-a) \left(\frac{wT^s - a}{w(1-a - a\Phi')}\right)^{1-a} \quad \text{for } \lambda \geq 0, \lambda \neq 1
\]  

(4.53)

\[
v(w) = \frac{(wT^s - a)\alpha^a (1-\alpha)^{1-a} (-\Phi')^a}{(1-a - a\Phi')w^{l-a}} \quad \text{for } \lambda = 1
\]  

(4.53')
The absolute and relative welfare losses are now given by the following expressions:

\[
ACWL = \frac{(\bar{w}^* + \bar{w})}{\left(\bar{w}^* + w\right)} \left[ \left( \frac{w - a}{\bar{w}^* + w} \right) \left( \frac{w^{\prime \prime} - a}{\left(1 - \alpha - \alpha \Phi'\right)} \right)^{\frac{1}{1 - \alpha}} \right] \quad \text{for } \lambda > 0, \lambda \neq 1 \quad (4.54)
\]

\[
ACWL = \left(\frac{\bar{w}}{w}\right)^{\frac{1}{1 - \alpha}} \left( \frac{w - a}{\bar{w}^*} \right) \left( \frac{w^{\prime \prime} - a}{\left(1 - \alpha - \alpha \Phi'\right)} \right)^{\frac{1}{1 - \alpha}} \quad \text{for } \lambda = 1 \quad (4.54')
\]

\[
RCWL = 1 - \left( \frac{w T^s - a}{w - a} \right) \left( \frac{\gamma w^* + w}{\gamma w^{\prime \prime} + w} \right) \left( \frac{a}{a^{\prime \prime}} \right)^{\frac{1}{1 - \alpha}} \quad \text{for } \lambda > 0, \lambda \neq 1 \quad (4.55)
\]

\[
RCWL = 1 - \left( \frac{w T^s - a}{w - a} \right) \left( \frac{\gamma w^* + w}{\gamma w^{\prime \prime} + w} \right) \left( \frac{a}{a^{\prime \prime}} \right)^{\frac{1}{1 - \alpha}} \quad \text{for } \lambda = 1 \quad (4.55')
\]

These expressions for ACWL and RCWL are functions of the slope of the efficient time budget constraint, which is a function of actual recuperation time (which can be found by solving expression (4.22) in the particular illness case) and thus of the wage rate. We can therefore only derive exact expressions for the absolute and relative welfare losses for the individuals with wages within the boundaries of this wage group numerically. This is done in example 2 in section 4.5.2.

The expressions for absolute and relative welfare loss for the low-wage individuals in case 2, regime 2, (indicated by the subscript 2,2 below) can be derived based on equations (4.28), (4.54') and (4.55') in the special cases of \(x = 0\) and \(x = 1\) for the C-D case:

\[
ACWL_{2,2} = \begin{cases} 
\bar{w}^{1-\alpha} a^\alpha (1-s), & x = 0 \\
\bar{w}^{1-\alpha} a^\alpha (r(1-\alpha(1-s)) - \alpha(1-r)\sigma'(1) - (sr - (1-r)\sigma'(1))^\alpha (sr)^{1-\alpha}) \left( (1-r)(1-\alpha + \alpha \sigma'(1)) - \alpha s r \right)^{\frac{1}{1-\alpha}} (1-\alpha)^{1-\alpha}, & x = 1
\end{cases}
\]  

(4.56)

\[
RCWL_{2,2} = \begin{cases} 
l, & x = 0 \\
l - \frac{(sr)^{1-\alpha} (sr - \sigma'(1)(1-r))^\alpha}{r(1-\alpha(1-s)) - (1-r)\alpha \sigma'(1)}, & x = 1
\end{cases}
\]  

(4.57)

Comparing the expressions for ACWL when wage is low (here: \(x = 0\)) and wage is higher (\(x = 1\)) given in expression (4.56), we find that the ACWL decreases when moving from the lower to the higher wage if the following condition holds:
The relative welfare loss decreases when moving from the lower to the higher wage if the following condition is fulfilled:

\[
1 - s s^a > \frac{r(1-\alpha(1-s)) - \alpha(1-r)\sigma(1) - (sr - (1-r)\sigma'(1))^a \left(\frac{sr}{(1-r)}\right)^{1-a}}{(1-r)(1-\alpha + \alpha\sigma(1)) - \alpha s r \left(\frac{sr}{(1-r)}\right)^{1-a}}
\]

This condition is fulfilled in illness case 2. The numerator of the fraction on the RHS is positive due to the assumption defining this illness case, i.e.: \(sr/(1-r) \geq \sigma'(0) \geq \sigma'(1)\), and the denominator is also positive (an alternative way of writing the denominator is the following: \(\frac{w-a}{a} (r(1-\alpha + \alpha\sigma'(1)) - \alpha s r)\)).

Hence, we see that whether ACWL decreases or increases when moving from a lower to a higher wage depends on the characteristics of the illness, i.e. its duration, \(r\), its severity, \(l-s\), and the marginal productivity in the post-recuperation period of additional recuperation, \(\sigma'(s)\), as well as the taste of the individual (i.e. the exponentials in the utility function, \(\alpha\)).

RCWL, on the other hand, decreases in wage.

Locating at the kink, where \(L = s r\), will only be chosen when wage lies above the wage, \(\bar{w}\), and below the critical wage, \(\bar{w}\), (based on equations (4.28) and (4.10)). The analysis is the same as under illness case 1, only the lower boundary wage rate differs. In the Cobb-Douglas case the absolute welfare losses decrease with increasing wages if the ratio of labour in the healthy scenario to labour in the sickness scenario is smaller than the marginal rate of substitution between efficient leisure and efficient labour in the sickness scenario to the power of \((1-\alpha)\), i.e. the weight given to leisure in the utility function, i.e.:

\[
\frac{\alpha(w-a)+a}{w(1-r)} < \left(\frac{\alpha s r}{(1-\alpha)(w(1-r)-a)}\right)^{1-a}
\]

We know from the analysis in illness case 1 that this condition is not fulfilled at the critical wage. When substituting in the lower boundary kink-wage given in expression (4.28), we find that the condition will hold if the following inequality holds:

\[
\left(1 + \alpha\sigma'(1) + \frac{r\alpha(1-s)}{1-r}\right) < \left(\frac{sr}{sr - (1-r)\sigma'(1)}\right)^{1-a}
\]

Relative welfare loss was found to be decreasing with increasing wages.

When comparing the welfare losses for the low-wage individuals in the illness case 2 with those predicted by the exogenous illness model (ACWLx) we find once again
that at the critical wage-rate in the exogenous-recuperation case the ACWL is higher than at the lower wage-boundary, whereas in illness case 1 the ACWL may be lower or higher but it is certainly less than the ACWL at the upper wage-boundary for the exogenous-recuperation case.

4.4.4 Interim Summary

In this section we have shown that in illness case 1, where due to the nature of the illness full recovery is imposed before returning to work, absolute comparable welfare loss will decrease with increasing wages until a certain wage rate between the existence wage and the critical wage, and will then start increasing again. Relative welfare loss will decrease with increasing wages throughout. In illness case 2, where low-wage individuals choose to work part of their recuperation time, we found that absolute welfare losses may be either increasing or decreasing with increasing wages up to the lower boundary kink-wage, depending on the characteristics of the illness and the taste of the individual. For those wage-groups that choose to exactly fully recuperate, absolute welfare loss may decrease or increase in wages, but if it is decreasing it will only do so up to a certain wage below the critical wage and will then start increasing again. Relative income loss will here again decrease with increasing wages throughout.

Hence, the individuals with the lowest wages may loose more welfare absolutely than individuals with higher wages up to a certain wage level due to the removal or restriction on the opportunity to earn additional income during illness. Relative welfare loss due to illness will be lower the higher the wage rate of the individual considered. The exogenous recuperation model (EXOR), on the other hand, predicts that the absolute welfare losses increase in wages throughout, whereas the relative income losses decrease in wages. This traditional way of modelling illness will therefore once again produce seriously misleading measurements, in particular for the lower wage-groups, if illness is a time that can be used in other activities. For those individuals who choose to recuperate fully, i.e. the higher-wage ones, the EXOR model performs better but the losses estimated for this wage-group using the EXOR model are higher by a constant factor because the individuals can derive no utility out of their sickness time in this model. The welfare loss predictions from the two models have been contrasted in figure 4.8 for illness case 1 and figure 4.9 for illness case 2.
In the illness case 1 the loss of opportunity to earn additional income during the illness period is total, whereas in the second illness case it is only partial; individuals can choose to work during illness but at a lower productivity and at the cost of an extended illness period. Therefore, the adverse health effect is worse in the first illness case, as supported by the income loss measurement. However, as already pointed out, consumption is only a part of welfare, and disregarding the effect of illness on leisure will be more misleading in illness case 1 where the opportunity to work more efficient hours has been removed and remaining time is taken out in leisure due to lack of an alternative option, than in illness case 2 where the individuals have a real choice between consumption and leisure. This can be illustrated by comparing the absolute welfare loss expressions with those of absolute income loss at lower wage-rates in illness cases 1 and 2. From table 4.5, where a numerical example of illness case is compared to an example of illness case 2, we observe that the ratio of the income loss in illness case 1 to the income loss in illness case 2 is smaller than the welfare loss ratio, implying that merely taking income losses into account would under-estimate the difference in adverse effects between the two illness cases.
Figure 4. 8: Welfare losses (absolute and relative) at various wage rates for illness case 1 and the EXOR model.

Figure 4. 9: Welfare losses (absolute and relative) at various wage rates for illness case 2 and the EXOR model.

Note: The figures above are based on the numerical examples given in section 4.5 of this chapter ($r = 0.3$, $\alpha = 0.5$, $\rho = 1$. In addition we have assumed the productivity during illness, $s$, in the case 1 to be 0.5 and in the case 2 to be 0.8.
4.5 Adverse Effects from Illness – Two Examples

In this section two numerical examples are presented illustrating the losses experienced in the two illness cases. We calculate the absolute and relative income and welfare losses, as well as the efficient time losses and work-absence predicted by the endogenous recuperation model and compare the results with those obtained from the exogenous model. We focus on the results for absolute and relative income and welfare losses, as the former are of interest in their own right and the latter is the correct measurement for the adverse effects of illness.

For the purpose of the two examples we assume a Cobb-Douglas preference structure with equal weights on consumption and leisure (i.e. \( \alpha = 0.5 \) and \( 1-\alpha = 0.5 \)), a net minimum consumption requirement identical for all individuals and equal to unity (\( a=1 \)), and a reference wage equal to 2 (\( w = 2 \)).

Since the net minimum consumption requirement and total healthy time are assumed to be equal to unity, the relevant wage rates must also be at least equal to unity in the healthy case (\( w \geq 1 \)). In table 4.1 below the amount of total efficient leisure, \( L \), efficient labour, \( H \), efficient time, \( T \), income, \( Y \), and the full income equivalent of indirect utility at a reference wage, \( F \), in the healthy scenario are given for some chosen wage rates.

Table 4.1: Healthy scenario \((a = 1, \bar{w} = 2, \alpha = 0.5)\)

<table>
<thead>
<tr>
<th>( W )</th>
<th>( L )</th>
<th>( H )</th>
<th>( T )</th>
<th>( Y )</th>
<th>( F )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>2.00</td>
<td>0.25</td>
<td>0.75</td>
<td>1.00</td>
<td>1.50</td>
<td>1.00</td>
</tr>
<tr>
<td>3.00</td>
<td>0.33</td>
<td>0.67</td>
<td>1.00</td>
<td>2.00</td>
<td>1.63</td>
</tr>
<tr>
<td>4.00</td>
<td>0.38</td>
<td>0.63</td>
<td>1.00</td>
<td>2.50</td>
<td>2.12</td>
</tr>
<tr>
<td>5.00</td>
<td>0.40</td>
<td>0.60</td>
<td>1.00</td>
<td>3.00</td>
<td>2.53</td>
</tr>
<tr>
<td>6.00</td>
<td>0.42</td>
<td>0.58</td>
<td>1.00</td>
<td>3.50</td>
<td>2.89</td>
</tr>
<tr>
<td>7.00</td>
<td>0.43</td>
<td>0.57</td>
<td>1.00</td>
<td>4.00</td>
<td>3.21</td>
</tr>
<tr>
<td>8.00</td>
<td>0.44</td>
<td>0.56</td>
<td>1.00</td>
<td>4.50</td>
<td>3.50</td>
</tr>
<tr>
<td>9.00</td>
<td>0.44</td>
<td>0.56</td>
<td>1.00</td>
<td>5.00</td>
<td>3.77</td>
</tr>
<tr>
<td>10.00</td>
<td>0.45</td>
<td>0.55</td>
<td>1.00</td>
<td>5.50</td>
<td>4.02</td>
</tr>
<tr>
<td>100.00</td>
<td>0.50</td>
<td>0.51</td>
<td>1.00</td>
<td>50.50</td>
<td>14.00</td>
</tr>
</tbody>
</table>

Source: Own calculations.

Looking at column 2 in table 4.1 we observe that the amount of leisure taken increases with increasing wages, but at a decreasing rate, e.g. increasing the wage by one
unit when the original wage is equal to 1 will increase the amount of leisure taken by 0.25, whereas increasing it by one unit when the original wage is 3 only increases the amount of leisure by 0.05. The reason for this behaviour is that with increasing wage rates the net minimum consumption requirement plays a decreasing role as a constraint upon the amount of leisure time that can be taken. Income is increasing at a constant rate with increasing wage rates due to the fact that full choice income (which increases by the same amount as the wage rate) is split equally between work and leisure. The full income equivalent of utility, on the other hand, is increasing at a decreasing rate because the wage rate affects the available choice time at a decreasing rate (with increasing wage rates the net minimum consumption requirement plays a decreasing role as a constraint upon choice time), and it only affects the amount of consumption derived from the efficient labour share of this choice time, not the amount of leisure derived from the efficient leisure share. Total efficient time does not change with the wage rate.

4.5.1 Example 1

In addition to the assumptions made in the healthy case, we now make some assumptions about the sickness parameters. In particular, let us assume that $\beta=0.8$, $r=0.3$, and $s=0.5$. In this case $\sigma'(0)=0.927$, $\sigma'(1)=0.464$, and $s/r/(1-r)=0.214$ which implies that we are dealing with illness case 1 (i.e. $s/r/(1-r) < \sigma'(1)$).

The critical wage was given in expression (4.10) and by substituting in the assumed sizes of the parameters implies that $\bar{w}=1.82$. At any wage rate below the critical wage and above the existence wage (here: $w \geq w_{min} = a/(1-r) = 1.43$) the individuals choose to exactly recuperate.

The tables 4.2 and 4.3 below give the size of the relevant variables and loss measurements for different wage rates in the case where recuperation time is endogenous (ENDOR model) and exogenous (EXOR-model), respectively. The amounts of efficient leisure, efficient labour, and total efficient time are given in columns 2, 3, and 4, respectively. Column 5 illustrates the level of income in the illness scenario, and column 6 the full income equivalent of utility at a reference wage. Column 7 shows the efficient time loss (ETL), whereas column 8 and 9 give the absolute income loss (AIL) and relative income loss (RIL), respectively, and columns 10 and 11 give the absolute welfare loss (ACWL) and relative welfare loss (RCWL). Absence from work (WA) due to illness is calculated in column 12.
Notice that since we assumed the net minimum consumption requirement (NMC) to be 1, the absolute losses in the tables below can be thought of as units of NMC, e.g. an income loss due to illness of 0.2 implies a loss of one-fifth of NMC-equivalents. Relative losses are percentage losses relative to the size of the variable when healthy.

When we use the endogenous recuperation model we see that the absolute income loss is larger the lower the wage rate (since $\alpha < 1 - r$) below the critical wage, and larger the higher the wage rate when above it, e.g. the lowest wage group ($w = 1.5$) will loose as much income as those with a wage rate of approximately 2.5 ($y^H - y^S = 0.20$). A similar finding is made for absolute welfare loss, where for the same wage rates the welfare loss is 0.33. The wage rate at which the welfare losses start increasing again lies below the critical wage, at a wage rate of 1.72. As with absolute income losses, the relative income losses decrease up to the critical wage rate and then start increasing again. However, it is only at very high wage rates that these losses are of a size comparable to those of the poorest individuals, e.g. individuals with a wage rate of 10 and of 1.6 have the same relative income loss ($y^H - y^S = 0.14$). We can finally observe that relative welfare loss is decreasing in wage throughout and the efficient time loss is the same across wages.

If the endogeneity of recuperation is not taken into account, then the absolute and relative income losses, as well as the absolute welfare losses, are larger the higher the wage, whereas the relative welfare loss is larger for the lower wage groups. The efficient time loss is the same across wages.

The figures 4.10 and 4.11 illustrate the losses from illness with endogenous and exogenous recuperation, respectively.
Table 4.2: Illness case 1 ($\beta=0.8$, $r=0.3$, $s=0.5$).

<table>
<thead>
<tr>
<th>$w$</th>
<th>L</th>
<th>H</th>
<th>T</th>
<th>Y</th>
<th>F</th>
<th>ETL</th>
<th>AIL</th>
<th>RIL</th>
<th>ACWL</th>
<th>RCWL</th>
<th>WA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.50</td>
<td>0.15</td>
<td>0.70</td>
<td>0.85</td>
<td>1.05</td>
<td>0.24</td>
<td>0.15</td>
<td>0.20</td>
<td>0.16</td>
<td>0.33</td>
<td>0.58</td>
<td>0.25</td>
</tr>
<tr>
<td>1.60</td>
<td>0.15</td>
<td>0.70</td>
<td>0.85</td>
<td>1.12</td>
<td>0.38</td>
<td>0.15</td>
<td>0.18</td>
<td>0.14</td>
<td>0.29</td>
<td>0.43</td>
<td>0.24</td>
</tr>
<tr>
<td>1.70</td>
<td>0.15</td>
<td>0.70</td>
<td>0.85</td>
<td>1.19</td>
<td>0.48</td>
<td>0.15</td>
<td>0.16</td>
<td>0.12</td>
<td>0.28</td>
<td>0.37</td>
<td>0.24</td>
</tr>
<tr>
<td>1.80</td>
<td>0.15</td>
<td>0.70</td>
<td>0.85</td>
<td>1.26</td>
<td>0.56</td>
<td>0.15</td>
<td>0.14</td>
<td>0.10</td>
<td>0.28</td>
<td>0.34</td>
<td>0.23</td>
</tr>
<tr>
<td>2.00</td>
<td>0.18</td>
<td>0.68</td>
<td>0.85</td>
<td>1.36</td>
<td>0.70</td>
<td>0.15</td>
<td>0.15</td>
<td>0.10</td>
<td>0.30</td>
<td>0.30</td>
<td>0.23</td>
</tr>
<tr>
<td>3.00</td>
<td>0.26</td>
<td>0.59</td>
<td>0.85</td>
<td>1.77</td>
<td>1.27</td>
<td>0.15</td>
<td>0.23</td>
<td>0.11</td>
<td>0.37</td>
<td>0.23</td>
<td>0.20</td>
</tr>
<tr>
<td>4.00</td>
<td>0.30</td>
<td>0.55</td>
<td>0.85</td>
<td>2.20</td>
<td>1.70</td>
<td>0.15</td>
<td>0.30</td>
<td>0.12</td>
<td>0.42</td>
<td>0.20</td>
<td>0.19</td>
</tr>
<tr>
<td>5.00</td>
<td>0.33</td>
<td>0.53</td>
<td>0.85</td>
<td>2.65</td>
<td>2.06</td>
<td>0.15</td>
<td>0.38</td>
<td>0.13</td>
<td>0.47</td>
<td>0.19</td>
<td>0.18</td>
</tr>
<tr>
<td>6.00</td>
<td>0.34</td>
<td>0.51</td>
<td>0.85</td>
<td>3.06</td>
<td>2.37</td>
<td>0.15</td>
<td>0.45</td>
<td>0.13</td>
<td>0.52</td>
<td>0.18</td>
<td>0.18</td>
</tr>
<tr>
<td>7.00</td>
<td>0.35</td>
<td>0.50</td>
<td>0.85</td>
<td>3.50</td>
<td>2.65</td>
<td>0.15</td>
<td>0.53</td>
<td>0.13</td>
<td>0.56</td>
<td>0.18</td>
<td>0.17</td>
</tr>
<tr>
<td>8.00</td>
<td>0.36</td>
<td>0.49</td>
<td>0.85</td>
<td>3.92</td>
<td>2.90</td>
<td>0.15</td>
<td>0.60</td>
<td>0.13</td>
<td>0.60</td>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td>9.00</td>
<td>0.37</td>
<td>0.48</td>
<td>0.85</td>
<td>4.32</td>
<td>3.13</td>
<td>0.15</td>
<td>0.68</td>
<td>0.14</td>
<td>0.64</td>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td>10.00</td>
<td>0.38</td>
<td>0.48</td>
<td>0.85</td>
<td>4.80</td>
<td>3.35</td>
<td>0.15</td>
<td>0.75</td>
<td>0.14</td>
<td>0.67</td>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td>100.00</td>
<td>0.42</td>
<td>0.43</td>
<td>0.85</td>
<td>43.00</td>
<td>11.88</td>
<td>0.15</td>
<td>7.50</td>
<td>0.15</td>
<td>2.12</td>
<td>0.15</td>
<td>0.15</td>
</tr>
</tbody>
</table>

Source: Own calculations.
Table 4.3: Sickness scenario with exogenous recuperation ($r=0.3$).

<table>
<thead>
<tr>
<th>w</th>
<th>L</th>
<th>H</th>
<th>T</th>
<th>Y</th>
<th>F</th>
<th>ETL</th>
<th>AIL</th>
<th>RIL</th>
<th>ACWL</th>
<th>RCWL</th>
<th>WA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5</td>
<td>0.02</td>
<td>0.70</td>
<td>0.70</td>
<td>1.05</td>
<td>0.06</td>
<td>0.30</td>
<td>0.20</td>
<td>0.16</td>
<td>0.52</td>
<td>0.90</td>
<td>0.25</td>
</tr>
<tr>
<td>1.6</td>
<td>0.04</td>
<td>0.70</td>
<td>0.70</td>
<td>1.12</td>
<td>0.13</td>
<td>0.30</td>
<td>0.18</td>
<td>0.14</td>
<td>0.54</td>
<td>0.80</td>
<td>0.24</td>
</tr>
<tr>
<td>1.7</td>
<td>0.06</td>
<td>0.70</td>
<td>0.70</td>
<td>1.19</td>
<td>0.21</td>
<td>0.30</td>
<td>0.16</td>
<td>0.12</td>
<td>0.55</td>
<td>0.73</td>
<td>0.24</td>
</tr>
<tr>
<td>1.8</td>
<td>0.07</td>
<td>0.70</td>
<td>0.70</td>
<td>1.26</td>
<td>0.27</td>
<td>0.30</td>
<td>0.14</td>
<td>0.10</td>
<td>0.57</td>
<td>0.68</td>
<td>0.23</td>
</tr>
<tr>
<td>2.0</td>
<td>0.10</td>
<td>0.60</td>
<td>0.70</td>
<td>1.20</td>
<td>0.40</td>
<td>0.30</td>
<td>0.30</td>
<td>0.20</td>
<td>0.60</td>
<td>0.60</td>
<td>0.23</td>
</tr>
<tr>
<td>3.0</td>
<td>0.18</td>
<td>0.52</td>
<td>0.70</td>
<td>1.55</td>
<td>0.90</td>
<td>0.30</td>
<td>0.45</td>
<td>0.23</td>
<td>0.73</td>
<td>0.45</td>
<td>0.20</td>
</tr>
<tr>
<td>4.0</td>
<td>0.23</td>
<td>0.48</td>
<td>0.70</td>
<td>1.90</td>
<td>1.27</td>
<td>0.30</td>
<td>0.60</td>
<td>0.24</td>
<td>0.85</td>
<td>0.40</td>
<td>0.19</td>
</tr>
<tr>
<td>5.0</td>
<td>0.25</td>
<td>0.45</td>
<td>0.70</td>
<td>2.25</td>
<td>1.58</td>
<td>0.30</td>
<td>0.75</td>
<td>0.25</td>
<td>0.95</td>
<td>0.38</td>
<td>0.18</td>
</tr>
<tr>
<td>6.0</td>
<td>0.27</td>
<td>0.43</td>
<td>0.70</td>
<td>2.60</td>
<td>1.85</td>
<td>0.30</td>
<td>0.90</td>
<td>0.26</td>
<td>1.04</td>
<td>0.36</td>
<td>0.18</td>
</tr>
<tr>
<td>7.0</td>
<td>0.28</td>
<td>0.42</td>
<td>0.70</td>
<td>2.95</td>
<td>2.08</td>
<td>0.30</td>
<td>1.05</td>
<td>0.26</td>
<td>1.12</td>
<td>0.35</td>
<td>0.17</td>
</tr>
<tr>
<td>8.0</td>
<td>0.29</td>
<td>0.41</td>
<td>0.70</td>
<td>3.30</td>
<td>2.30</td>
<td>0.30</td>
<td>1.20</td>
<td>0.27</td>
<td>1.20</td>
<td>0.34</td>
<td>0.17</td>
</tr>
<tr>
<td>9.0</td>
<td>0.29</td>
<td>0.41</td>
<td>0.70</td>
<td>3.65</td>
<td>2.50</td>
<td>0.30</td>
<td>1.35</td>
<td>0.27</td>
<td>1.27</td>
<td>0.34</td>
<td>0.17</td>
</tr>
<tr>
<td>10.0</td>
<td>0.30</td>
<td>0.40</td>
<td>0.70</td>
<td>4.00</td>
<td>2.68</td>
<td>0.30</td>
<td>1.50</td>
<td>0.27</td>
<td>1.34</td>
<td>0.33</td>
<td>0.17</td>
</tr>
<tr>
<td>100</td>
<td>0.35</td>
<td>0.36</td>
<td>0.70</td>
<td>35.50</td>
<td>9.76</td>
<td>0.30</td>
<td>15.00</td>
<td>0.30</td>
<td>4.24</td>
<td>0.30</td>
<td>0.15</td>
</tr>
</tbody>
</table>

Source: Own calculations.
Figure 4. 10: Loss from illness – case 1 (\(\beta=0.8, r=0.3, s=0.5\))

![Graph showing loss from illness case 1 with wage rate on x-axis and numeric loss on y-axis.]

Figure 4. 11: Loss from illness with exogenous recuperation (\(r=0.3\)).

![Graph showing loss from illness with exogenous recuperation with wage rate on x-axis and numeric loss on y-axis.]

Source: Own calculations and illustrations
4.5.2 Example 2

Let us assume that $\beta=0.5$, $r=0.3$, $s=0.8$. In this case $\sigma'(0)=0.225$, $\sigma'(1)=0.18$, and $sr/(1-r)-0.343$ which implies that we are dealing with illness case 2 (i.e. $sr/(1-r)>\sigma'(0)$).

The critical wage determining the lower wage-boundary for the group of individuals who choose not to work during illness can once more be found by substituting the relevant parameters into expression (4.10) and is found to be $\bar{w}=2.17$. The poorest individuals have a wage between the existence wage (here: $w \geq w_{\min} = a/s = 1.25$) and the lower boundary kink-wage (here: $w \leq \bar{w} = a/((1-r)(\sigma'(1)+1)-sr) = 1.71$). There is, finally, a group of people with wages higher than 1.71 but lower than 2.17 ($1.71<w<2.17$) who will locate at the kink.

Table 4.4 below gives the size of the relevant variables and loss measurements for different wage rates in the case where recuperation time is endogenous (ENDOR model). For the exogenous recuperation model (EXOR) the findings will be identical to those in table 4.3 for the same wage rates. Columns 3, 4, 5, 6 and 7 give the amounts of efficient leisure, efficient labour, efficient time, income, and the full income equivalent of utility at a reference wage, respectively. Column 2 shows the amount of the full recuperation time actually spent recuperating, $x$. Efficient time loss is shown in column 8, whereas column 9 and 10 give the absolute and relative income loss, respectively, and columns 11 and 12 the absolute and relative welfare loss. Work-absence is calculated in Column 13.

Relative welfare losses decrease in wages throughout, whereas absolute welfare losses decrease with increasing wages up to a certain wage rate below the critical wage rate (here: 2.05) and then start increasing again. However, it is only at relatively high wage rates that the absolute welfare losses are of a similar size to those of the lower wage groups, e.g. to experience the same welfare loss (absolutely) as those individuals with a wage rate of 1.5 ($P^H-P^S = 0.21$) we have to go to a wage rate of 6. Absolute and relative income losses, on the other hand, increase with increasing wages up to the kink, then decrease for increasing wages at the kink, and finally increase with increasing wages for those individuals who can afford post-recuperation leisure. An interesting finding is that in this particular example the poorest individuals with a survival wage of 1.25 loose as much in absolute income ($P^H-P^S = 0.13$) as those with a wage of approximately 1.85 and 4. For relative income loss, on the other hand, no individual with a wage higher than approximately 1.75 loose as much as the poorest individuals considered. Finally, we observe that the efficient time loss is 20% for the poorest individuals ($w=1.25$) and only 6% (less than a third) for any wage group with a wage rate higher than 1.8.
If the endogeneity of recuperation is not taken into account (table 4.6), then the results are similar to those found under example 1.

Figure 4.12 illustrates the losses from illness with endogenous recuperation. In the case of the EXOR-model, the loss picture will be identical to figure 4.11.
Table 4.4: Illness case 2 ($\beta=0.5$, $r=0.3$, $s=0.8$).

<table>
<thead>
<tr>
<th>w</th>
<th>x</th>
<th>L</th>
<th>H</th>
<th>T</th>
<th>Y</th>
<th>F</th>
<th>ETL</th>
<th>AIL</th>
<th>RIL</th>
<th>ACWL</th>
<th>RCWL</th>
<th>WA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.25</td>
<td>0</td>
<td>0.00</td>
<td>0.80</td>
<td>0.80</td>
<td>1.00</td>
<td>0.00</td>
<td>0.20</td>
<td>0.13</td>
<td>0.11</td>
<td>0.32</td>
<td>1.00</td>
<td>-0.73</td>
</tr>
<tr>
<td>1.3</td>
<td>0.17</td>
<td>0.04</td>
<td>0.79</td>
<td>0.83</td>
<td>1.02</td>
<td>0.09</td>
<td>0.17</td>
<td>0.13</td>
<td>0.11</td>
<td>0.29</td>
<td>0.78</td>
<td>-0.56</td>
</tr>
<tr>
<td>1.4</td>
<td>0.45</td>
<td>0.11</td>
<td>0.76</td>
<td>0.87</td>
<td>1.06</td>
<td>0.24</td>
<td>0.13</td>
<td>0.14</td>
<td>0.12</td>
<td>0.25</td>
<td>0.51</td>
<td>-0.29</td>
</tr>
<tr>
<td>1.5</td>
<td>0.66</td>
<td>0.16</td>
<td>0.74</td>
<td>0.90</td>
<td>1.11</td>
<td>0.37</td>
<td>0.10</td>
<td>0.15</td>
<td>0.12</td>
<td>0.21</td>
<td>0.37</td>
<td>-0.09</td>
</tr>
<tr>
<td>1.6</td>
<td>0.84</td>
<td>0.20</td>
<td>0.72</td>
<td>0.92</td>
<td>1.15</td>
<td>0.49</td>
<td>0.08</td>
<td>0.15</td>
<td>0.12</td>
<td>0.18</td>
<td>0.27</td>
<td>0.09</td>
</tr>
<tr>
<td>1.7</td>
<td>0.99</td>
<td>0.24</td>
<td>0.70</td>
<td>0.94</td>
<td>1.19</td>
<td>0.60</td>
<td>0.06</td>
<td>0.16</td>
<td>0.12</td>
<td>0.16</td>
<td>0.20</td>
<td>0.23</td>
</tr>
<tr>
<td>1.8</td>
<td>1</td>
<td>0.24</td>
<td>0.7</td>
<td>0.94</td>
<td>1.26</td>
<td>0.71</td>
<td>0.06</td>
<td>0.14</td>
<td>0.10</td>
<td>0.14</td>
<td>0.16</td>
<td>0.23</td>
</tr>
<tr>
<td>1.9</td>
<td>1</td>
<td>0.24</td>
<td>0.7</td>
<td>0.94</td>
<td>1.33</td>
<td>0.80</td>
<td>0.06</td>
<td>0.12</td>
<td>0.08</td>
<td>0.13</td>
<td>0.14</td>
<td>0.23</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0.24</td>
<td>0.7</td>
<td>0.94</td>
<td>1.40</td>
<td>0.88</td>
<td>0.06</td>
<td>0.10</td>
<td>0.07</td>
<td>0.12</td>
<td>0.12</td>
<td>0.23</td>
</tr>
<tr>
<td>2.1</td>
<td>1</td>
<td>0.24</td>
<td>0.7</td>
<td>0.94</td>
<td>1.47</td>
<td>0.95</td>
<td>0.06</td>
<td>0.08</td>
<td>0.05</td>
<td>0.12</td>
<td>0.12</td>
<td>0.22</td>
</tr>
<tr>
<td>2.5</td>
<td>1</td>
<td>0.27</td>
<td>0.67</td>
<td>0.94</td>
<td>1.68</td>
<td>1.21</td>
<td>0.06</td>
<td>0.08</td>
<td>0.04</td>
<td>0.13</td>
<td>0.10</td>
<td>0.21</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0.30</td>
<td>0.64</td>
<td>0.94</td>
<td>1.91</td>
<td>1.49</td>
<td>0.06</td>
<td>0.09</td>
<td>0.05</td>
<td>0.15</td>
<td>0.09</td>
<td>0.20</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>0.35</td>
<td>0.60</td>
<td>0.94</td>
<td>2.38</td>
<td>1.95</td>
<td>0.06</td>
<td>0.12</td>
<td>0.05</td>
<td>0.17</td>
<td>0.08</td>
<td>0.19</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>0.37</td>
<td>0.57</td>
<td>0.94</td>
<td>2.85</td>
<td>2.34</td>
<td>0.06</td>
<td>0.15</td>
<td>0.05</td>
<td>0.19</td>
<td>0.08</td>
<td>0.18</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>0.39</td>
<td>0.55</td>
<td>0.94</td>
<td>3.32</td>
<td>2.68</td>
<td>0.06</td>
<td>0.18</td>
<td>0.05</td>
<td>0.21</td>
<td>0.07</td>
<td>0.18</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>0.40</td>
<td>0.54</td>
<td>0.94</td>
<td>3.79</td>
<td>2.98</td>
<td>0.06</td>
<td>0.21</td>
<td>0.05</td>
<td>0.22</td>
<td>0.07</td>
<td>0.17</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>0.41</td>
<td>0.53</td>
<td>0.94</td>
<td>4.26</td>
<td>3.26</td>
<td>0.06</td>
<td>0.24</td>
<td>0.05</td>
<td>0.24</td>
<td>0.07</td>
<td>0.17</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>0.41</td>
<td>0.53</td>
<td>0.94</td>
<td>4.73</td>
<td>3.52</td>
<td>0.06</td>
<td>0.27</td>
<td>0.05</td>
<td>0.25</td>
<td>0.07</td>
<td>0.17</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>0.42</td>
<td>0.52</td>
<td>0.94</td>
<td>5.20</td>
<td>3.76</td>
<td>0.06</td>
<td>0.30</td>
<td>0.05</td>
<td>0.27</td>
<td>0.07</td>
<td>0.17</td>
</tr>
<tr>
<td>100</td>
<td>1</td>
<td>0.47</td>
<td>0.48</td>
<td>0.94</td>
<td>47.50</td>
<td>13.15</td>
<td>0.06</td>
<td>3.00</td>
<td>0.06</td>
<td>0.85</td>
<td>0.06</td>
<td>0.15</td>
</tr>
</tbody>
</table>

Source: Own calculations.
Figure 4.12: Loss from illness – case 2 (β=0.5, r=0.3, s=0.8).

Source: Own calculations and illustration.
4.5.3 Comparison of Income and Welfare Losses in the Two Examples

In example 1 we have analysed a relatively severe illness lasting 30% of a given time-period and during which productivity is only 50% of what it would be under healthy conditions. This illness is furthermore characterised by a relatively linear post-recuperation productivity function (the parameter measuring the curvature of the productivity, which is a function of the amount of actual recuperation, is assumed to be: $\beta = 0.8$). The less severe illness case in example 2 also lasts 30% of the time period, but is characterised by a productivity of 80% and a post-recuperation productivity function with a larger curvature ($\beta=0.5$). Any amount of actual recuperation larger than 0 but less than full recuperation gives a higher amount of efficient labour with a larger curvature of the post-recuperation productivity function.

The absolute and relative income and welfare losses are higher at each wage rate in the more severe illness case. For the low wage rates this is due to the fact that in the illness case 2 the opportunity to earn additional income has only been partially removed. In the case of the higher wage-groups it is due to the individuals drawing less efficient leisure time out of the illness period when they are more severely ill. The only exception to the above finding is found for those wage rates with which individuals under both illness scenarios end up at the kink of the efficient time budget constraint - for these wages the income loss is clearly the same (see figure 4.13). The “true” adverse effect measurement, welfare loss, is however still higher in the worst illness case for these wages, capturing the fact that in illness case 1 they have no choice but to locate at the kink (no more efficient labour to be gained by taking less leisure), whereas in illness case 2 it is a choice.

Looking at the ordinal ranking of the adverse effects due to illness for the low-wage groups we see that in illness case 1 income loss performs well compared to the ranking given by welfare loss. The lowest wage group ($w = 1.5$) will lose as much both in absolute income and absolute welfare as those with a wage rate of approximately 2.5. Both types of losses decrease in increasing wages up to a certain wage rate, and then start increasing in wages. The same comparison fares less well in the second illness case. First of all, for the lowest wage-groups the absolute income loss is increasing in wages, whereas the absolute welfare loss is decreasing. Second, the welfare loss of the lowest wage-group is only matched by the loss of those with a wage above 10, whereas the absolute income loss of the lowest wage-groups is also experienced by those with a wage-rate of about 1.85 and 4.5. The worse ranking performance of income loss in the second
illness case is due to the fact that the individuals with a wage above the existence wage could have obtained more efficient labour (and consumption), but prefer the choice they have made with less consumption and more leisure. In the more severe illness case, on the other hand, they would prefer to have more consumption but cannot (the leisure is imposed and does not give much utility as long as the minimum consumption requirement is barely fulfilled).

Looking at the absolute performance of income loss compared to welfare loss, a different picture appears. In table 4.5 we give the ratio of absolute income loss to absolute comparable welfare loss for various wage rates in the two illness cases. In the case of the lower wage-groups, income loss performs distinctly worse in the first illness case than the second, implying that income loss is a better proxy for welfare loss when the choice is merely restricted than when it is non-existent. Although the income loss is larger in the first illness case than the second, the welfare loss ratio is even larger. The reason for this is that since we are approaching the existence wage in illness case 1 the individuals do not derive significant utility out of their imposed leisure time, whereas in illness case 2 we are still considering wage-rates with which the individual chooses to take some time in leisure (i.e. recuperation). Omission of the lost welfare effect of leisure matters more in the former.

For any wage rate that in both illness cases lies above the critical wage rate we observe that the ratio is the same in both illness cases. This merely reflects the fact that these wage-groups face a time of choice which they can allocate freely between consumption and leisure, and that in both illness cases they are assumed to have the same preferences. The ratio is unity at a wage rate of 8 due to the specific parameter specifications in these examples.40

40 AIL above the critical wage rate is given by: AIL = αwr(1-s), and ACWL by ACWL = (\bar{w} / w)^2 wr(1 - s). When α=0.5, \bar{w} =2, and w=8 the two expressions are the same.
Table 4.5: Comparisons of income and welfare losses in cases 1 and 2.

<table>
<thead>
<tr>
<th>W</th>
<th>AIL/ACWL</th>
<th>AIL/ACWL</th>
<th>AIL 1/</th>
<th>ACWL 1/</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CASE 1</td>
<td>CASE 2</td>
<td>AIL 2</td>
<td>ACWL 2</td>
</tr>
<tr>
<td>1.5</td>
<td>0.60</td>
<td>0.68</td>
<td>1.43</td>
<td>1.62</td>
</tr>
<tr>
<td>1.6</td>
<td>0.62</td>
<td>0.83</td>
<td>1.20</td>
<td>1.61</td>
</tr>
<tr>
<td>1.7</td>
<td>0.57</td>
<td>1.04</td>
<td>1.00</td>
<td>1.83</td>
</tr>
<tr>
<td>1.8</td>
<td>0.49</td>
<td>1.02</td>
<td>1.00</td>
<td>2.08</td>
</tr>
<tr>
<td>2.0</td>
<td>0.50</td>
<td>0.81</td>
<td>1.50</td>
<td>2.43</td>
</tr>
<tr>
<td>3.0</td>
<td>0.61</td>
<td>0.61</td>
<td>2.50</td>
<td>2.50</td>
</tr>
<tr>
<td>4.0</td>
<td>0.71</td>
<td>0.71</td>
<td>2.50</td>
<td>2.50</td>
</tr>
<tr>
<td>5.0</td>
<td>0.79</td>
<td>0.79</td>
<td>2.50</td>
<td>2.50</td>
</tr>
<tr>
<td>6.0</td>
<td>0.87</td>
<td>0.87</td>
<td>2.50</td>
<td>2.50</td>
</tr>
<tr>
<td>7.0</td>
<td>0.94</td>
<td>0.94</td>
<td>2.50</td>
<td>2.50</td>
</tr>
<tr>
<td>8.0</td>
<td>1.00</td>
<td>1.00</td>
<td>2.50</td>
<td>2.50</td>
</tr>
<tr>
<td>9.0</td>
<td>1.06</td>
<td>1.06</td>
<td>2.50</td>
<td>2.50</td>
</tr>
<tr>
<td>10.0</td>
<td>1.12</td>
<td>1.12</td>
<td>2.50</td>
<td>2.50</td>
</tr>
</tbody>
</table>

Source: Own calculations.

Figure 4.13: Income loss versus welfare loss for illness cases 1 and 2.
4.5.4 Relevance

The World Bank publishes data on the percentage of the population in different countries below a nationally or internationally defined poverty line. Looking at e.g. the international poverty line defined at $2 a day (1985 international prices, adjusted for purchasing power parity), we see that a country such as the Arab Republic of Egypt has 51.9 per cent of its population situated below that poverty line, whereas in the case of India this percentage is 88.8.

If we think of the poverty line as being slightly higher than the minimum consumption requirement in our model, here assumed to be equal to one, (thinking of it as higher than the minimum consumption requirement means that it is possible to fall below it) some interesting insights ensue. Recall that a wage rate of 1 is necessary in order for a healthy individual to exactly fulfil the net minimum consumption requirement when working full time, and that under illness this survival wage rate is higher than 1 (1.25 in the less severe illness case, and 1.43 in the more severe illness case). Let us for simplicity suggest that somewhere below a wage rate of 2 lies the poverty line (the exact location depends on the sickness specifications) in the model that has been presented. This implies that the range of wage rates within which both absolute and relative welfare losses, and to a lesser degree absolute and relative income losses are decreasing in wages would encompass 52 and 89 per cent of the Egyptian and Indian population, respectively.

---

41 World Bank (1998), p. 64.
4.6 Summary

This chapter has investigated the adverse effects of illness for different wage groups, based on the endogenous recuperation model described and set out in chapters 2 and 3. In particular, it has focused on the implications of using various adverse effect measurements. One relatively severe illness case (case 1) and one less severe illness case (case 2) have been discussed and compared to the traditional exogenous-recuperation model.

Due to the illness characteristics, working during the full recuperation time in illness case 1 does not yield additional efficient labour time, whereas in illness case 2 trade-off between efficient leisure and efficient labour is possible during the full recuperation period. The minimum wage-rate required in order to survive is therefore lower in the latter, implying that poorer people are able to survive when the illness is less serious because they are able to earn additional income during their illness.

Absence from work due to illness has been a commonly used statistical measurement of illness. We have seen that if there is a tendency for low-wage individuals to work although being ill in order to earn enough to sustain a minimum level of consumption, then the health effect for these wage-groups would not appear, or at least be underestimated, using work-absence as adverse effect measurement. Furthermore, this illness measurement will not distinguish imposed full recuperation from illness before returning to work (due to the severity of the illness) from chosen recuperation.

Efficient time loss can be considered one possible measurement of the adverse health effect of illness that avoids the problem of not capturing the illness of people who work (or would like to work) while being sick. Efficient time loss is a measure of the amount of time a person is ill weighted by the productivity during illness (whether in work or leisure). However, although this measurement includes the productivity loss experienced by individuals who carry on working, as well as the productivity loss in the production of leisure time, it does not take into account that the efficient time spent in leisure and labour may not be worth equally much to the individual. Indeed, under certain circumstances an individual with less efficient time during an illness may be better off than an individual with more efficient time – the determining factor here is the minimum consumption requirement.
In illness case 2 low-wage individuals may choose to work part of their recuperation time and it is shown that this may extend the time period when their productivity is low. The efficient time loss would hence be larger for this wage-group than for the higher-wage ones whereas the recorded work-absence would be lower. Work-absence therefore performs particularly badly in illness case 2. In illness case 1, the efficient-time-loss would record equal illness effects for all wage groups, whereas work-absence would increase with lower wages due to the fact that they would have worked more during that period had they been healthy than their richer counterparts. Hence, the performance of the two illness measurements has in this case reversed, with efficient time loss being distinctly worse.

The measurement of income loss differs in two main ways from the measurements discussed above. Since workers are paid according to productivity, income loss due to illness would record the adverse effect on both those individuals who stay at home and on those who go to work, in contrast to work-absence. Furthermore, whereas the measurement of efficient time loss implicitly counts leisure time and labour time as equally valuable, income loss looks only and explicitly at efficient work time loss (weighted by the wage rate).

We have seen that under certain conditions the lowest-wage individuals may loose more income, both relatively and absolutely, than higher wage individuals due to the same illness (identical illness characteristics). We have further shown that the income losses are worse in illness case 1, where the loss of opportunity to earn additional income during the illness period is total, than in illness case 2 where it is only partial - individuals can choose to work during illness but at a lower productivity and at the cost of an extended illness period. Income losses start increasing in wages when the critical wage is reached (the wage rate where individuals choose to recuperate fully), due to the fact that preferences over choice consumption and leisure now determine the amounts chosen. Efficient work time loss is now essentially the same for all within this wage group (may vary if the elasticity of substitution between consumption and leisure is different from 1) and the wage rate therefore determines the difference in income loss. Nevertheless, it is only at very high wage rates that absolute and relative income losses are the same as for the poorest individuals. Relative income loss, which is identical to the relative loss in efficient working hours, is higher for higher wages due to the fact that the total amount of
healthy working hours to which we compare the hours lost is lower for these wage-groups.

However, income loss as a measurement of the overall effect of illness is still unsatisfactory. Consumption is only a part of welfare, and when used to calculate the monetary value of health-impacts, this Cost-Of-Illness (COI) approach has been criticised for ignoring other welfare aspects of being healthy (e.g. the effect on leisure time).

Disregarding the effect of illness on leisure will be more misleading in illness case 1, where the opportunity to work more efficient hours has been removed and remaining time is taken out in leisure due to lack of an alternative option, than in the case where the individuals have a real choice between consumption and leisure, i.e. illness case 2. This is found when we compare income loss to welfare loss due to illness. The measurement of welfare loss is the only measurement reviewed in this chapter that considers the full adverse effect of illness taking individual preferences into account – it is hence the proper adverse effect measurement against which the other measurements are judged. The full income expenditure function is chosen to measure welfare here since it is the classic money metric of utility, which can be computed easily and directly compared with net income. Furthermore, the full income equivalent is measured at a reference wage (equal for all individuals), a procedure that makes welfare comparable over individuals.

Similarly to income losses, welfare losses have under certain conditions been found to be larger, both relatively and absolutely, for the lowest-wage individuals than for higher-wage ones. As was the case for income loss, this is partly due to the fact that the opportunity to earn additional income during illness has been partially or wholly removed. In addition, the closer the wage-rate of the individual is to the existence wage the less utility he derives from leisure. Absolute welfare loss will decrease with increasing wages until a certain wage rate between the existence wage and the critical wage is reached. At that wage rate the measured losses will start increasing again because the amount of income is high enough so that the imposed leisure time yields some utility. Relative welfare loss will continue decreasing in wages throughout, the reason being that welfare under illness is compared to lower initial welfare in the case of the poorer, and the minimum consumption requirement on wage earnings therefore matters more.

We have seen in this chapter that whatever measurement of adverse effects from illness we use, the traditional way of modelling illness as time lost, i.e. recuperation as exogenous, misrepresents the adverse effect of illness if illness is in fact a time of choice.
The errors that would result from using this model-approximation are particularly serious for the lower wage-groups, as they will choose to work during the recuperation time if possible. For those individuals who choose to recuperate fully, i.e. the higher-wage ones, the EXOR model performs better but the losses estimated for this wage-group using the EXOR model are higher by a constant factor because the individuals can derive no utility out of their sickness time in this model.

4.7 Conclusion

Little attention has previously been given in the literature to the choice an individual may face on how to spend his or her sickness time, nor has the effect of that choice upon the duration of sickness time been analysed. This analysis has attempted to cover this important gap.

Using standard utility maximisation theory we analyse labour supply behaviour under constraints imposed by sickness and minimum consumption requirements. In particular, the conditions under which an individual may decide to work rather than stay at home during the recuperation time from sickness are examined. Departing from the previous literature, productivity in work and utility derived from leisure (productivity of leisure time) in this model are influenced by the labour supply decision: if the individual works during exogenously imposed recuperation time, productivity in the post-recuperation time suffers, and produces a non-linear efficient time budget constraint. This new approach required a more advanced way of modelling sickness than what has previously been used. Sickness is in the presently developed model identified by three parameters: duration, severity in terms of productivity loss, and efficiency of recuperation.

The analysis finds that if any leisure is taken in the post-recuperation phase then recuperation is spent entirely in leisure (recuperating). The reason for this is that recuperating during the recuperation phase has the additional advantage of increasing the productivity in both work and leisure in the post-recuperation phase. Hence, when the individual chooses to take out the equivalent of the recuperation time or less in leisure she will only take it during the recuperation time. An implication of the above finding is that the more leisure an individual takes, the more likely she is to stay at home during recuperation.
For a rather general utility function, an empirically supported elasticity of substitution between leisure and consumption, and under the assumption that no sick pay is received, we therefore find that only higher-wage individuals will choose to recuperate fully, whereas others will work while being sick.

Main implication of this finding is that, under certain conditions, the income and welfare losses (relative and absolute) from illness are larger for the lowest wage groups than for those with higher wages. This, in turn, suggests that the occurrence of certain common illnesses or illness-reducing policies (for instance vaccination-projects or pollution-reduction initiatives) may have income and welfare redistributing effects. The exogenous recuperation model, on the other hand, predicts absolute income and welfare losses to uniformly increase in wages and may hence be seriously misleading, especially for the lower wage groups.

Furthermore, commonly-used measurements of illness can lead to distorted estimates of illness by neglecting the interaction between income and behavioural choices. Four types of illness-measurements and the implication of their use were presented.

Work-absence is shown to be a poor measurement of the adverse effects of illness because it does not capture the illness of those who work although being sick, and does not distinguish between imposed and chosen recuperation. It is particularly misleading in the case of less-severe illnesses where there are efficient labour hours to be gained by working while sick.

Efficient time loss is not well-suited as a measurement of illness effects as it does not take into account that the efficient time spent in leisure and labour may not be worth equally much to the individual. This measurement is especially ill suited for use in the case of more severe illnesses, where the individual would have preferred to have less overall efficient time but more efficient labour time had it been possible.

Income loss as an adverse effect measurement captures the fact that in illness case 1 the loss of opportunity to earn additional income during the illness period is total, whereas in illness case 2 it is only partial. Income loss records the adverse effect of illness on those who go to work as well as those who stay at home during the illness, in contrast to work-absence. Whereas the measurement of efficient time loss counts leisure time and labour time as equally valuable, income loss looks only and explicitly at efficient work time loss (weighted by the wage rate) which matters more for the low-wage groups.
However, the income loss measurement ignores other welfare aspects of being healthy (e.g. the effect on leisure time).

Welfare loss is the correct measurement of the adverse health effect as it takes individual's preferences for leisure and consumption into account. By measuring the full (choice) income equivalent at a reference wage we have made the losses comparable over individuals.

An interesting question (and an issue that deserves further analysis) looks at the implications of introducing a sick pay policy into our model. We will limit our analysis to the issues of full sick pay and poor-only sick pay, and will assume that there are no problems of moral hazard and no possibility of obtaining any sick pay for time spent working during part of the recuperation period. Under the concept of 'full sick pay' we understand receiving full compensation while being off sick for those hours one would usually work when healthy. Under the concept 'poor-only sick pay' we mean that sick pay (in this case full sick pay) is only made to those who have such a low wage rate that, given the net minimum consumption requirement on their wages they would choose to work during recuperation if that implies more efficient working time (i.e. they have a wage rate at or below the critical wage in our model).

First, on the positive side we will briefly discuss how the analysis of recuperation behaviour would be affected by the introduction of sick pay. With full sick pay, everybody will recover fully and even the poorest individuals who barely survive even under healthy conditions (and who would not survive without sick pay when falling ill) will now also survive under illness. With poor-only sick pay the same will be true, since all the non-poor choose to recuperate fully anyway even without sick pay. As for income, no income losses are incurred with full sick pay, whereas with poor-only sick pay only the non-poor will incur income losses. For the non-poor, income losses will increase with their wages.

Interestingly, with full sick pay the low-wage individuals will be better off (in terms of welfare) under sickness than in the healthy scenario, whereas a high-wage individual who cares mostly about leisure would still suffer welfare losses. The reason for this perverse effect of illness for the poor is that under healthy conditions the low-wage individuals spend most of their time working. If they were to get sick, they would continue receiving the full-productivity wage as handouts, while being in fact at home recuperating, from which they would draw some leisure utility. On the other hand, high-
wage individuals with a high preference for leisure activities will still suffer under illness because they get less utility out of their leisure time when ill. Nevertheless, they would be better off with than without sick pay.\(^{42}\)

At an amount of sick pay less than the full amount, the low-wage individuals will have to trade off the amount they would receive staying home with the one they could earn while continuing working, taking into account the fact that they may actually work more hours during the sickness period than they would have had they been healthy during the same time-period. Depending on the size of the parameters it may still be an optimal choice for the poorest to work during the recuperation time.

The next question relates to the issue of who pays the sick pay; individual firms or society. If the firm pays sick pay in our model, it would basically pay the same amount as under healthy conditions but for less production. For everybody who would survive even without sick pay during illness, compensation will imply lower production than under a system of no compensation. The intuition here is quite clear: in the model without sick pay, wage is paid according to efficiency and any person who chooses to work during illness does so because it increases his or her production and hence income. Only in the case of the lowest-wage individuals, i.e. those who would otherwise not survive, would sick pay lead to higher production, but by paying these workers sick pay the firm would be paying them more than their marginal product. A small, price-taking, profit-maximising firm would therefore not choose to pay sick pay in the current model.

A proper analysis of benefit packages and contracts, especially in an inter-temporal framework, would have to take into account a host of other issues such as how firms go about attracting their preferred types of employees, the psychological effect on production of having some sort of health insurance, the uncertainty over sickness developments if employees do not recover fully from their illness, training costs if additional staff have to be taken on, internalisation of production externalities if sickness is related to the workplace etc.

Regarding society, as opposed to the firm, we can assume it is maximising a social welfare function. As we have seen in our analysis of absolute welfare and income losses without sick pay, it is possible that the lowest-wage individuals may incur larger losses than the higher-wage individuals from a certain illness. With full sick pay for all no income losses are incurred, however the low-wage individuals may in fact experience a

\(^{42}\)This analysis is partial and disregards how sick pay is paid for.
welfare gain from falling sick under this compensation-scheme, whereas the higher-wage individuals may still suffer. Although a proper analysis is required to understand the welfare effects of introducing sick pay, taking both the welfare effects on different wage-groups and their willingness or ability to pay compensation into account, there seems to exist the possibility of potential Pareto improvement, whether it is a compensation for all or a poor-only sick pay policy.

Nevertheless, paying sick pay to everybody, irrespective of the wage rate, will lead to a deadweight loss because the high-wage individuals were going to recuperate anyway. The deadweight loss is here related to the fact that with full sick pay, individuals can substitute some of the work in the post recuperation period for leisure and hence productivity and output will decrease. In the severe illness case, sick pay will allow some of those low-wage individuals who could not survive without the policy to actually survive. For less severe illness cases, sick pay will prevent them from working during illness.

In addition to the above-mentioned deadweight loss, losses could result from the inefficiencies inherent in operating a sick pay program: e.g. administrative costs, the cost of moral hazard through increased reported illness in response to the existence of this program, and any deadweight loss from financing these benefit payments.

Finally, we will try to answer a normative question, namely; would it be desirable to have sick pay? We have seen that income and welfare losses from a given illness may be larger for the low-wage individuals (including death in the case of the very poorest) than for those with higher wages, and hence having serious equity effects. For this reason, sick pay would clearly be a redistributive policy increasing equity, especially when paid only to the low-wage individuals, and may be desirable for this reason. A more compelling argument for centrally provided sick pay or public intervention is inefficiency or simple absence of credit and insurance markets, which prevents an individual to protect against dire effects of illness.

Another interesting theoretical extension of the model would be to consider the joint determination of recuperation time and other health investments (e.g. medical expenditures, obtaining medical treatment). Finally, the relevance and importance of the insights from the model developed in this first part of the thesis will become clearer once the model is tested empirically.
PART 2: WHY LEVEL AND DISTRIBUTION OF INCOME MATTER FOR THE HEALTH COSTS OF AIR POLLUTION
Chapter 5: Why Level and Distribution of Income Matter for the Health Costs of Air Pollution

5.1 Introduction

Numerous empirical studies have observed associations between air pollution and various adverse health events. The information retrieved from such studies has been used to estimate the health impact of air pollution and quantify the potential health benefits of controlling air pollution. To date, the possible effect that income may have on actual as well as measured health impact has been largely ignored in these studies.

This chapter will question the correctness of the way in which the health impact of air pollution has been quantified and valued in industrialised countries. In addition, we point out that even for correct estimates from industrialised country studies, the effects of income and other personal characteristics may pose significant problems for transferability into developing country contexts.

The chapter explores four ways in which the level and distribution of income of a population exposed to air pollution may have an impact on the quantification and valuation of the adverse health effects associated with air pollution. First, if the adverse health effects from air pollution depend on personal characteristics and individual income level, then failure to include these variables in the regression analysis will lead to the effect estimates being biased (omitted variable bias). Second, even with unbiased estimates for one particular population sample, transferring an adverse health effect estimate onto another population with different characteristics will be incorrect (transfer challenge). Third, even with unbiased estimates for one particular population sample, the quantification of the health effect in that population will be incorrect given that different income groups have different willingness to pay to avoid illness (covariance challenge). Finally, the traditional measurements of the health effect from air pollution may be particularly unsuited to capture the true effect on the poorer income groups due to differing behavioural responses to illness in different income groups (information bias). Similarly, the degree to which ambient pollution levels can be used as a measurement of exposure to air pollution may vary with income. We will suggest corrections for or improvements to the shortcomings.
The next section reviews briefly how, in the literature, the health impact of air pollution has been valued and quantified using epidemiological dose-response studies and willingness-to-pay (WTP) studies. The dose-response function is then analysed within the wider framework of health production functions in section 5.3, and modelling clarifications necessary for our further discussion are made. Section 5.4 discusses the various reasons why the measured health effects from air pollution may vary between income groups, whereas section 5.5 looks at the effect income may have on the willingness-to-pay. Section 5.6 considers the problems posed by income and income distribution both to the quantification and cost estimation of the health impact of air pollution, and to the transfer of estimates from one population to another, and makes suggestions for improvements. The last section summarises the main insights and concludes with some policy implications.

Note that although factors such as season, humidity, temperature, and pollutant mixture may all affect the estimated health impact of air pollution, they are not the focus of this chapter and will not be discussed further. The issues of the existence of threshold levels of pollution (below which no effect can be detected), as well as the shape of the dose-response relationship have received some attention in the dose-response literature (EPA volume III (1996), Holgate et al. (1999), Schwartz (1994a)) but will not be dealt with in this discussion.

5.2. Health Impact of Air Pollution – Quantification, Valuation, and Application

The main method used to establish whether there is a relationship between air pollution and the occurrence of certain illnesses is the use of epidemiological studies. These studies investigate the statistical relationship between the concentration of air pollutants in a particular area, and the occurrence of particular health effects in a sample of the population, and they are known as dose-response-functions. Significant dose-response relationships for several of the main air pollutants have been shown (e.g. Ostro et al. (1996), Simpson et al. (1997), Schwartz and Dockery (1992), Dockery et al. (1992), Krupnick and Cropper (1989), Cropper et al. (1997), Anderson et al. (1996), Bacharova et al. (1996), Pope (1991), Pönkä (1990), Rossi (1993) and Schwartz et al. (1993)). Health impacts include eye irritation, respiratory diseases, cardiovascular effects and premature

43 Strictly speaking, the commonly used term "dose-response function" only applies when actual doses to target organs are used in the analysis. However, in the following the term 'dose-response' will usually refer to the effect of ambient concentration of pollution on health and the term 'exposure-response' will refer to
deaths. Evidence further suggests that the association between air pollution and adverse health effects is a causal one (e.g. Seaton et al. (1995), Dockery et al. (1993), Maddison et al. (1996)).

The epidemiological investigations involve three main study designs: statistical inference based on cross-sectional data, time-series data, and cohort studies. The first compares the rate or prevalence of specific health events across a sample of locations at a given point in time. The second examines changes in a health outcome in one specific location as air pollution levels change over time. The third follows a selected population sample (e.g. from different cities) over time. The two latter methods using time-series are usually preferred because they have the advantage of reducing the problem associated with confounding or omitted variables.

Measurement of human exposure to air pollution or dose to the target organs is usually based on data obtained from centrally located outdoor monitoring stations, although lately some studies have also included measurements of indoor air pollution in order to calculate the actual exposure more exactly. As for changes in health, most epidemiological studies report the relative increase in the risk of experiencing an effect when going from one concentration level to a higher one, expressed as the relative risk ratio, \( RR = \frac{p_1}{p_0} \). \( p_1 \) and \( p_0 \) are the average frequencies (prevalence or incidence) in the population at the two different concentration levels. Mortality is thus measured as the change in the probability of dying, whereas morbidity can be measured in a variety of ways. Some studies classify morbidity by reported cases of a disease or of a symptom, some measure illness by symptom days (i.e. take duration of symptom into account), one classification is according to whether the illness experienced is chronic or acute, and finally many studies classify morbidity by degree of impairment of activity (e.g. restricted activity days, bed disability days).

Most epidemiological studies to date have concentrated on the acute mortality impact of air pollution, as the phenomenon only requires the examination of the risks posed by contemporaneous pollution levels and is thus easy to implement and insensitive to confounders. However, this approach has been questioned on the grounds that only the vulnerable part of the population (mostly the old and the very sick) seem to be exposed to the pollution to which the individual is exposed on adverse health.

---

44 See Lipfert and Wyzga (1997) for a more detailed discussion on exposure measurement.

45 See Cropper and Freeman III (1991) for a further discussion on morbidity measurements.

46 Although in recent years a couple of prospective cohort studies of chronic mortality from particulate air pollution have been published (e.g. Dockery et al. (1993), Pope et al. (1996)).
affected by air pollution to the point of dying. The life span reduction (also known as years of life lost – YLL) is probably small compared to the average human life span implicit in value of life studies. Studies of death due to chronic exposure require the additional analysis of lagged response to past pollution levels stretching back over a number of years, while controlling for individual characteristics, or alternatively of cross sectional studies. Chronic mortality impacts of air pollution may well be more important than the acute effects upon which most cost benefit analyses are based.47

If individual characteristics $Z_i$ affect the health effect $H_i$ an individual $i$ experiences from a certain increase in air pollution concentration $P_j$, then the typical dose-response function may be written in the following way:

$$H_h = \sum_{i=1}^{N} H_{h}(P_j, Z_i)$$

(5.1)

Here, $N$ represents the population exposed to pollutant $j$, and $h$ represents the type of adverse health event (e.g. number of cases of cardiovascular deaths, number of asthma attacks). The estimated dose-response relationship in the relevant population for a given set of individual characteristics is then given by:

$$b_{hy} = \frac{dH_{h}}{dP_j}$$

(5.2)

where $b_{hy}$ is the slope of the dose-response function of health effect $h$ per year with respect to pollutant $j$ holding individual characteristics fixed.

The health change due to a change in the ambient level of a certain air pollutant, calculated using the dose-response procedure, is multiplied by a price in order to calculate the cost of the health impact (see figure 5.1). Several approaches to assigning a “price” (monetary value) to the health impact have been proposed in the literature. Freeman (1979) categorises these approaches as determining values according to individual preferences (willingness to pay), according to resources or opportunity costs (also known as the cost-of-illness approach (COI)), or by political processes. He argued, however, that the latter two are inferior to the willingness to pay (WTP) approach. The political process approach is based on the assumption that governments act purposefully, rationally, and consistently, so that the willingness to pay implicitly observed in the governmental decisions is the correct one. The main problem with using the political-process-approach is the assumption of consistency and rationality of the political process, especially in the

face of information gaps concerning the effects of different policies on health. The COI-
approach incorporates decreases in an individual’s expected life-time earnings as well as
personal and social expenditure due to this individual’s illness. The approach is based on
two strong assumptions; first, that earnings reflect the individual’s marginal productivity
(i.e. his or her contribution to total economic output), and second, that productivity
changes can be used to value loss of welfare. A major problem with assuming that
earnings reflect productivity is the omission of non-market productivity. More
importantly, this approach ignores the direct welfare aspect of being healthy and is
therefore not the theoretically correct measure of the benefits enjoyed by people from
improved air quality. Some measurement of WTP to reduce the probability of falling ill
or of dying derived either from direct questioning/stated preferences (contingent
valuation), from market-based studies (e.g. labour market studies) or from the averting
behaviour method is therefore considered conceptually superior to the other approaches.
An individual’s willingness to pay $V_H$ for a change in the probability of an adverse health
effect occurring $dH$, is the income (more generally wealth) that must be taken away from
her, $dY$, to compensate her for the reduction in probability while keeping utility $v$
constant:

$$V_H = -\frac{dV}{dv} \frac{dH}{dY}$$  \hspace{1cm} (5.3)

Since one person’s wellbeing or health may also enter another person’s preference
function, the WTP to avoid a risk or an illness suffered by a particular person should
ideally be the sum of his or her own WTP and that of other people who care about her.
The latter may, however, be difficult to obtain empirically. In addition to the sum of
WTPs, the illness imposes costs on society not taken into consideration in the WTP
estimates (i.e. externalities) and these should therefore be added to the WTPs to give the
monetary value of the health impact. The main objection put forward against the WTP
method is that when it is used as a welfare measure, rich individuals will get a larger
weight than poor people (due to their larger budgets). In other words, an additional pound
in the pocket of a rich person causes the same welfare gain as it would in the pocket of a
poor person.
The estimated dose-response relationship and the associated price estimate have been used to calculate the costs or benefits of changes in air pollution levels in the studied population. In addition, these estimates have been transferred to other populations, with or without amendments, as a substitute for an actual study in order to estimate the benefits of pollution regulation in these populations. Until recently all quantification of the health effects from air pollution in developing countries was based on transfers of dose-response functions from developed countries, mainly the U.S. Further developing country studies are required, although the quality of the data at the moment is a major constraint.

Income and its potential influences on the above described estimation procedures have barely been discussed in the literature so far. A few exceptions, however, exist. There have been some empirical studies of the income elasticity of WTP for reducing the risk of adverse health effects and these have usually found the income elasticity of WTP to be less than 1 (see section 5). However, the theory behind the effect of income on WTP appears to have been little discussed. Furthermore, Lave and Seskin (1971) have regressed mortality on both air pollution, as well as socio-economic factors, such as income and age, and found that income plays a significant role for the mortality rate for some specific diseases. The coefficient for income when regressing mortality rate on pollution and income, however, only captures the independent effect of income on mortality (e.g. poor people will tend to live closer together, something which enhances the risk of catching Tuberculosis - a communicable disease). How income affects the association between exposure and mortality (and thus the dose-response function) is, however, not explained by such a regression analysis.

---

48 Tuberculosis is for example a disease of the poor, whereas cancer is more prevalent among people with higher incomes.
5.3 The Model

We will in the following set out the model used in the subsequent discussions, which originates from the health-production literature, and show the relationship to the dose-response function. The focus is on the role of individual characteristics and individual behaviour. Furthermore, in the latter part of this section we will show that different types of measured health events may to a varying extent allow for different individual responses. This implies that the measured coefficient of the dose-response function, and the effect of individual characteristics and income on this coefficient, may vary according to how the health event is captured. In particular, a distinction will be made between the probability of illness occurring and the actual illness event. This distinction has not previously been made in the theoretical health production function literature, and will be important to keep in mind when discussing the various effects of income on the estimated relationship between health and pollution.

5.3.1 Health Production Tradition

In a paper by Cropper and Freeman (1991) it was argued that the dose-response function could be considered a reduced form which can be derived from the health production function. In the health production function averting and mitigating activities are crucial to the effect of air pollution. Both averting and mitigating measures are types of health investment that have been thoroughly discussed in the "health-as-capital" literature (e.g. Grossman (1972), (1982), (2000), Mushkin (1962), Cropper (1977), Muurinen (1982), and Wagstaff (1986)). Averting activities reduce the exposure to air pollution, whereas mitigating activities reduce the health impact from pollution exposure. The costs of these activities may include time as well as out-of-pocket costs. Hence, pollution has not only a direct effect on health through exposure, but in addition affects health indirectly through its influence on mitigating and averting activities. Cropper and Freeman assumed that even when controlling for pollution, both averting and mitigating activities may vary with individual characteristics.

The individual characteristics of particular interest for our purposes are the baseline health level $B_i$, and income level $Y_i$. The adverse health experience (i.e. $A_i$) is determined by the individual's exposure to pollution $E_i$, and the person's ability $A_i$. Exposure is determined by the person's income level and the pollution levels in the area where the person lives.

---

49 Strictly speaking, income level is not an individual characteristic but a variable determined partly by individual characteristics (e.g. ability) and partly by factors exogenous to the individual (e.g. labour unions, competition). However, for simplicity of exposition we will assume that it is given and varies from one individual to the next – hence, it can be treated as an individual characteristic.
morbidity or mortality effect) of individual $i$, $H_i$, was modelled in Cropper and Freeman (1991) as a function of the level of mitigating activity $M_i$, as well as exposure to air pollution $E_i$, which again is a function of the ambient pollution level $P$ and the level of averting activities undertaken by the individual $A_i$.\footnote{It may be worth pointing out that the exposure that causes us to fall ill, and the exposure during the recuperative period may in reality differ since there is a certain time aspect involved. Illness duration is necessarily preceded by the occurrence of illness, but we will assume that the pollution level is unchanged between the time when the illness occurs and it develops. This implies that the changes in ambient pollution we are looking at in the dose-response function are longer-term changes. Since there is much uncertainty about the time-lag of detectable health effects, we suggest that this simplifying assumption should not cause much concern.} Note that we drop the subscripts below, as there is no reason for misunderstanding and it will leave the equations uncluttered. If we also allow the adverse health event to vary with the individual baseline health while holding exposure and mitigation fixed, then the health production function can be written in its generic form as:

$$H = h[E(A, P), M, B]$$

assume:

\begin{align*}
H_M &< 0, H_E > 0, H_B < 0, E_A < 0, E_P > 0
\end{align*} \hspace{1cm} (5.4)

In order to obtain the demand functions for averting and mitigating activity, we need to maximise utility subject to the budget constraint and the health production function. Let utility be a function of consumption $C$ and illness occurrence $H$:\footnote{For ease of exposition of the arguments in this section we exclude leisure as input into the utility function.}

$$U = f(C, H) \quad U_C > 0, U_H < 0$$ \hspace{1cm} (5.5)

Furthermore, income is spent on consumption, averting activity, and mitigating activity:\footnote{The price of consumption has been normalised to 1. Furthermore, for ease of exposition of the main insights in this section we assume that income is given, and not affected by the adverse health effects from air pollution.}

$$Y = C + p_M M + p_A A$$ \hspace{1cm} (5.6)

The demand for mitigating and averting activities are then both functions of the prices of mitigation and averting activity, $p_M$ and $p_A$ respectively, income, the level of pollution, and individual characteristics:

$$M = m(P, \bar{P}_M, \bar{P}_A, Y, B) \quad \text{with} \quad M_Y > 0, M_{p_M^A}^M < 0$$ \hspace{1cm} (5.7)

$$A = a(P, \bar{P}_A, \bar{P}_M, Y, B) \quad \text{with} \quad A_Y > 0, A_{p_A^A}^A < 0$$ \hspace{1cm} (5.8)

The effect of pollution, baseline health and knowledge on mitigating and averting activities cannot be signed based on the given information, nor can we say something about the cross price effects. By substituting expressions (5.7) and (5.8) for averting and mitigating activities in the (adverse) health production function (5.4), we can now take
the total derivative of the adverse health event with respect to pollution, which gives us the dose-response relationship:

\[
\frac{dH}{dP} = \frac{\partial H}{\partial E} \frac{dE}{dP} + \frac{\partial H}{\partial A} \frac{dA}{dP} + \frac{\partial H}{\partial M} \frac{dM}{dP} \quad H_p > 0 \tag{5.9}
\]

On the basis of the equations above, and assuming for the moment simple linear equations, we can now set out the fundamental regression model which will form the basis of much of our discussions in this paper. From equation (5.4) we can write adverse health event \( H \) (morbidity or mortality) as a function of exposure \( E \), mitigating activities \( M \), and baseline health \( B \).

\[
H = \beta_0 + \beta_1E - \beta_2M - \beta_3B + u \tag{5.10}
\]

\( u \) is a random error term and the \( \beta \)s are the coefficients to be estimated. Mitigating and averting activities are both functions of pollution \( P \), income \( Y \), baseline health level \( B \), and education \( K \), and are given in expressions (5.11) and (5.12) respectively:

\[
M = \alpha_0 + \alpha_1Y + \alpha_2P + \alpha_3B + \lambda \quad \text{assume } M_y > 0, M_p > 0 \tag{5.11}
\]

\[
A = \delta_0 + \delta_1Y + \delta_2P + \delta_3B + \eta \quad \text{assume } A_y > 0, A_p > 0 \tag{5.12}
\]

where \( \lambda \) and \( \eta \) are the random errors and the coefficients are denoted with \( \alpha \)s and \( \delta \)s. Finally, exposure is a function of the pollution level, averting activities, and random error \( \varepsilon \).

\[
E = \sigma_1P - \sigma_2A + \varepsilon \quad \text{assume } E_p > 0, E_A < 0 \tag{5.13}
\]

Substituting for \( M \), \( A \), and \( E \) from expressions (5.11), (5.12), (5.13) into equation (5.10), we obtain the reduced form of the model:

\[
H = \gamma_0 + \gamma_1P - \gamma_2Y - \gamma_3B + \nu \quad \text{where}
\begin{align*}
\gamma_0 &= \beta_0 - \beta_2\alpha_0 - \beta_3\sigma_2\delta_0 \\
\gamma_1 &= \beta_1 - \beta_2\alpha_2 - \beta_3\sigma_2\delta_2 \\
\gamma_2 &= \beta_4\alpha_1 + \beta_1\sigma_2\delta_1 \\
\gamma_3 &= \beta_4\alpha_3 + \beta_3 \\
\nu &= u - \beta_2\lambda - \beta_1\sigma_2\eta + \beta_3\varepsilon
\end{align*} \tag{5.14}
\]

Hence, the total derivative from expression (5.9) is equivalent to \( \gamma_1 \) in our regression model. The model set out above will suffice to explain three of the challenges posed by income that are discussed in this paper. However, the fourth challenge, the transfer challenge, is based on the idea that the model in fact is likely to be non-linear. In particular, individual characteristics may arguably influence the direct effect of a certain dose of exposure on health, the effect of a change in pollution on the amount of averting and mitigating activities undertaken, and the effect of the mitigating activities on health.
and of the averting activities on exposure. A certain dose-response relationship would then clearly be accurate only for populations with identical characteristics.

5.3.2 A Clarification: Why the Role of Income May Vary with Health Event Specification

Because we are particularly interested in the many ways, both direct and indirect, in which income affects the value of the health effects of pollution changes, it will become important to distinguish the type of health effect we are studying. There are two categories of health events that are usually estimated in the dose-response literature; one is the probability of falling ill or of dying, whereas the other also takes the severity of the illness into account (in terms of duration and/or impairment). The distinction is important to make as there is no reason why income should have the same effect on the probability as it has on severity. The main difference between these measurements is that mitigating and averting activities do not necessarily play the same role in the two.

Harrington and Portney (1987) distinguish between preventive and recuperative mitigation. In this paper we find it more useful to distinguish between mitigation, which implies recuperative mitigation, and baseline health which includes previous health investments.\(^{53,54}\)

Let us model loss of health from a particular illness of an individual as the product of the probability of falling ill with this illness \(q_h\), and the severity of the illness once it occurs, \(S_h\) (i.e. when \(q=1\)):

\[
H = S_h|_{q_h} \times q_h \quad (5.15a)
\]

We can also think of \(q_h\) as being the probability of dying from a certain illness, but we would clearly not multiply it with any measure of severity since it is implicit in the probability. If pollution causes the illness from which a person eventually dies, then we suggest treating the illness episode and death as two separate occurrences, where the loss of health from illness is modelled as above.\(^55\)

\(^{53}\) Given the amount of different health risks other than air pollution facing an individual, the amount of mitigating activities directly aimed at reducing the health effect of air pollution before they occur are likely to be negligible.

\(^{54}\) Averting activity can also in principle be subdivided into preventive and recuperative averting activities. The optimal level of averting activity may change with the onset of the illness, since new information upon which the decision is based has been added.

\(^{55}\) It may be useful to distinguish between two ways in which exposure to pollution can lead to death; in one case the individual suffers from some illness unrelated to pollution (e.g. possibly cardiovascular problems, Asthma) but dies of this illness due to air pollution, in the other case the exposure leads to the illness (e.g. possibly some types of cancer) and then eventually to death. In the latter case the probability of dying may
The probability of falling ill is here expressed as a function of exposure and baseline health:
\[ q = q(E, B) \]
Once the illness or symptom occurs, the level of mitigating activities undertaken by the individual, as well as exposure and baseline health determine the severity of the illness:
\[ S = S(M, E, B) \]
Combining the various functions, the health production function can be written:
\[ H = q_s(E, B) \times S_s(M, E, B) \quad (5.15b) \]
By taking health-production into account, it is therefore clear that the dose-response coefficients for the entire illness event, \( dH/dP \), and for the probability of occurrence, \( dq/dP \), are likely to differ, and that individual characteristics may affect different adverse health effect measurements differently.\(^{56}\) The effects of individual characteristics, and in particular income, on the estimated dose-response coefficients will also depend on the particular adverse health measurement used.\(^{57}\)

Three main implications of this probability versus severity/duration analysis are worth drawing attention to. First, if we are interested in the full health effect of air pollution (at given individual characteristics), looking at the probability of falling ill will be misleading as long as the level of pollution also affects the path/duration of the sickness (via its effect on exposure and/or on mitigating activities).\(^{58}\) Second, even if we were only interested in estimating the costs associated with an increased probability of falling ill (i.e. without taking into account the influence of pollution on duration and severity) problems are encountered when assigning a price to the changed probability. A willingness to pay to decrease the probability of falling ill will implicitly be based on an assumption about severity and duration and is hence the "wrong" price to be used for the changed probability of falling ill. Finally, when discussing how income may affect the

---

133

be affected by mitigating activities during the illness and averting behaviour, whereas in the former only averting activities will play a role.

56 For a mathematical derivation of the alternative dose-response coefficients refer to appendix 5.A.

57 For a mathematical derivation of the effect of income on the alternative dose-response coefficients refer to appendix 5.A.

58 Only in the case of mortality is the level of pollution after the onset of the adverse health effect irrelevant. However, if pollution first caused the illness from which the individual eventually dies, an interesting insight is obtained: severity may offset duration. After the individual has fallen ill, a high level of continued exposure may lead to increasing severity and bring the time of death forward, whereas with little exposure the person may be ill for a long time before dying. If illness is preferred to death, no matter the degree of severity, then we obtain the perverse effect of a lower WTP in the case of an illness of short duration than of long duration.
valuation process it will be important to keep in mind that this depends on the way we are measuring adverse health effects.

5.4 Why the Dose-Response Coefficient May be Affected by Income

If we accept the fact that sickness affects both wellbeing and productivity, and that individual behaviour may influence the amount of sickness experienced, then there are several paths through which one would expect income, and income-related characteristics, to affect the estimated association between ambient air pollution exposure and health. The first two paths discussed below are concerned with how income can influence the *actual* health effect of exposure, whereas the latter two have to do with the way health impacts and exposure are being measured:

i) mitigating measures
ii) baseline health
iii) averting activity and exposure
iv) behavioural choice and health effects

More generally, income has two roles to play in affecting the estimated dose-response coefficients; as an *effect modifier*, and as a *confounder*. In the former case income affects the relationship between two correlated variables, whereas in the latter income will bias the estimated relationship between two correlated variables if it is not controlled for in the regression analysis. Both of these roles will be discussed below when relevant.

5.4.1 Mitigating Measures

The optimal behavioural response to health changes may vary between different individuals, and may affect the development of an existing health effect. Although a sickness would under certain conditions be of a certain severity and length, these may be changed by own decisions – illness and recuperation from illness are hence endogenous. This type of investment in health, known as *mitigating* activity, is concerned with decreasing the *effect* (or the *health response*) of a certain dose of exposure rather than decreasing the dose itself.

Mitigating measures include seeking medical advice, taking medicines or vitamins, following a healthy diet, and taking time off from work to recuperate.\(^{59}\) These

---

\(^{59}\) In measuring the effect of staying away from work on the development of the illness caused by air pollution, one complication may be encountered: staying indoors may entail both lower continued exposure and a lower activity level. Unless the indoors exposure is measured it will therefore be difficult to separate
activities are a matter of choice, and the decision will depend on time, income, the existence of sick pay or health insurance, prices, information, risk perceptions, and uncertainty.

Assuming health is a normal good, and we spend money on mitigating activities (e.g. medicines, consultations, nutrition, health insurance) to obtain better health, mitigating expenditure will tend to increase with higher income levels. One would further expect the choice over labour supply under sickness in different income groups to differ. Particularly in developing countries, where the social welfare system often is insufficient, employees are not offered paid sick days. In addition, credit markets are usually not available for low-income households, especially not for consumption purposes, and furthermore low income tends to be associated with low savings. Due to these market imperfections or missing markets, poor people may have to work although they are sick in order to fulfil their minimum consumption requirements. The time constraint imposed by the minimum consumption requirement may also prevent low-wage individuals from seeking medical advice and treatment.

There is a line of research within the Grossman health-as-capital tradition that has been analysing the effect of uncertainty on health care spending by the poor. Dardanoni and Wagstaff (1986) suggest that greater absolute risk aversion of the poor (due to relatively little financial capital) makes them adjust their portfolios away from risky investments in health capital. Selden (1993), on the other hand, argues that the poor are the most vulnerable to any losses in income due to adverse health effects. Since the poor can less afford to be sick, they should be relatively more inclined to spend on preventive measures that would increase their 'precautionary' stocks of health capital. These arguments are relevant in the present context, as there is uncertainty involved in the development of an existing illness. This line of research relies on the poor having a certain room for choice, whereas the rest of the discussion on labour supply during illness and medical investments was based on severely constrained or non-existent choice due to low income, no credit, and no sick-pay. These may therefore best be viewed as arguments relevant at different levels of poverty, rather than as conflicting arguments.

A potential difference between low and high-income countries is the quality and availability of health care. All high-income countries, except for the United States, have

---

the effect of averting and mitigating behaviour in this case.
60 For a further discussion of the reasons why individuals may be working instead of recuperating when ill, and the implications for the illness-path, income, and welfare, please refer to chapter 2. A detailed analysis
achieved universal coverage, whereas low-income countries need to expand the coverage to the poor. Furthermore, apart from the government budget (which is limited by a narrow tax base and weak collection capacity), limited savings, the small size of the formal employment sector, weak institutions, and an underdeveloped financial sector limit the options for formal insurance schemes. Households therefore often rely on informal arrangements to provide protection in the case of illness. The implication of these facts is that even if an individual in a low-income country would like to undertake more mitigating activities, these may simply not be available. The quality of health care in developing countries is in addition often poor, which decreases the efficiency of mitigating measures (Schieber (1997)).

We argued in section 5.3 that health can be modelled as a function of both exposure and mitigating behaviour, and mitigating activity as a function of exposure and income (variables not relevant for the argument have been left out).\(^6\) We can write the relationships in a generic form:

\[ H = h(M(E, Y \ldots), E) \]

The cross-partial derivative and the partial derivative of the adverse health effect with respect to income are given by the following expressions:

\[
\frac{d^2 H}{dE dY} = \frac{\partial H}{\partial M} \frac{\partial^2 M}{\partial E \partial Y} + \frac{\partial^2 H}{\partial M^2} \frac{\partial M}{\partial E} \frac{\partial M}{\partial Y} + \frac{\partial^2 H}{\partial E \partial M} \frac{\partial M}{\partial Y}
\]

\[
\frac{\partial H}{\partial M} \frac{\partial M}{\partial Y}
\]

It is necessary to distinguish between changes/ variations in the budget constraint affecting behaviour and thereby the association between exposure and health (expression (a)), and those affecting health at a given level of exposure (expression (b)). In the former case, income affects the relationship between exposure and health even when it is linearly controlled for in the regression equation, and is known as a modifier variable or moderator. In the latter, income will bias the estimated health effect of air pollution exposure if it is not controlled for and if exposure is correlated with income, and it is known as a confounder variable. Note that the change in income we refer to could be either a change that occurs across time or it could be a cross-sectional variation. In section of the endogenous recuperation idea can be found in chapters 2 - 4.

\(^6\) We model mitigation as a function of exposure here, rather than of pollution as in equation (5.8) in section 5.3. It does not change the main idea of the argument and is algebraically more convenient. It is not clear which modelling form is more correct. If both mitigation and averting activity are functions of ambient pollution levels, this means that the optimal amounts of both activities are decided simultaneously. If the individual first decides on the amount of averting activity to undertake and thereupon, given the resulting exposure, decides on the amount of mitigating activities, then mitigation can be modelled as a function of exposure. Exposure, in turn, would be a function of pollution, averting activities, and the prices of averting and mitigating activities.
5.4.3 we will argue that both cross-sectional and time-series correlation between exposure and income may be a reality, and in section 5.6.1 we discuss the direction of bias for the estimated coefficient if income is omitted.

With the time, income, and institutional constraints discussed above we will expect to observe that the amount of mitigating activities undertaken for a certain amount of exposure increases in income, i.e. that expression (b) takes a negative sign. Some empirical findings may be interpreted as support to this idea. Alberini et al. (1997) found that expenditure on mitigating activities is increasing with income cross-sectionally. Furthermore, the global estimate of the income elasticity of health expenditure has been found to be 1.13. (Schieber (1997)).

Let us turn to the role of income as an effect modifier. It is likely that a change in exposure has a smaller adverse effect on health for larger levels of mitigating activities, because it would otherwise not make sense to invest in these. Furthermore, it is plausible that a change in exposure will lead to a larger change in the amount of mitigating activities undertaken for higher income groups. Although the effect of mitigation on health may be decreasing in mitigation, added together the overall effect of exposure on adverse health is likely to decrease in income, both cross-sectionally (within and between locations/cities) and across time, i.e. expression (a) above takes a negative sign.

To summarise, mitigating measures imply costs to the individual, either direct costs or opportunity costs in terms of lost time and income, which poor people may not be able to afford. On the other hand, if the poor have a choice they may also ill-afford to let these investments be. The association between air pollution exposure and adverse health effects may be influenced by own behaviour, and could be stronger for lower income groups.

5.4.2 Baseline Health

Baseline health level will tend to be associated with adverse health incidences independently of air pollution exposure. Unless baseline health and exposure are uncorrelated, failing to control for baseline health will bias the estimated health effect of air pollution, i.e. it will act as a confounder. We argue below that baseline health will tend to be lower for lower income groups, and in section 5.4.3 that exposure may be higher for

---

62 This would be the case if individuals choose the most efficient mitigating activities first.
lower income groups, so that baseline health and exposure may in fact be (negatively) correlated.

More importantly, however, the association between health effects, $H$, and exposure, $E$, may vary from one population group to the next because of differing baseline health, $B$, in other words, baseline health may act as an effect modifier. Note that baseline health does not only affect the association between exposure and health directly, but also indirectly by possibly affecting the amount of mitigating expenditure and its efficiency in reducing the adverse health effects.

Individuals with a weakened physical state, a history of diseases, as well as elderly individuals are thought to be more vulnerable to air pollution. It seems intuitively likely that poor health leads to a relatively stronger direct association between exposure and health \textit{ceteris paribus} (i.e. the first term on the right hand side of the equation will tend to be negative). Nevertheless, it is important to keep in mind that the indirect effect via mitigating activities may also differ between individuals or population groups with high and low baseline health. In particular, it is plausible that mitigating expenditures have a larger effect on health for individuals with low baseline health than for those with higher levels, and that due to being more vulnerable, people in a weakened physical state will invest relatively more in mitigation. On the other hand, elderly people and people with poor health have a shorter life expectancy, something which \textit{ceteris paribus} will tend to decrease investments in health. Findings by Alberini et al. (1997) and Krupnick et al. (2000) suggest that expenditure on mitigating activities increases with lower baseline health if the baseline health factors considered are chronic and prior respiratory illnesses.\footnote{Refer to section 5.6.2 for a more detailed discussion of these studies.} If low levels of baseline health lead to higher amounts of mitigation undertaken, then this may partly counteract the effect low baseline health has in increasing the association between exposure and health, however, it is unlikely that they will cancel out. Empirical dose-response studies have indeed found that mortality among the elderly is more responsive to changes in particulate pollution than is total mortality or mortality among the younger generation (Ostro et al. (1996), Schwartz and Dockery (1992)). Evidence further suggests that air pollution has its greatest adverse effects on people with pre-existing chronic conditions such as asthma, bronchitis, and emphysema (Ostro (1987)). To our knowledge, dose-response studies that take into account other
baseline health factors such as malnutrition or the presence of diseases affecting the immune system (e.g. aids) have so far not been undertaken.

We may expect the baseline health level in different income groups to differ. In particular, we would expect low-income groups to have a relatively low average health level, something which has strong empirical support (Doorslaer et al. (1997), Eachus et al. (1999), Michelozzi et al. (1999), Walker (2000)). Assuming low-income groups have relatively weak health, it is plausible that the adverse health effect from a certain change in the level of ambient pollution is larger in this group when age is controlled for. Taking age into account, one would expect the adverse health effect on low-income groups to occur earlier in life than for high-income groups, because at each age they will have lower baseline health. Indeed, studies carried out in developed countries tend to find a stronger effect of pollution on mortality in the group of the population over 65 years of age, whereas Cropper et al. (1997) only find a significant impact of pollution on death before the age of 65 in their study of pollution mortality in New Delhi.

There are a number of reasons why one would expect low-income people to have worse initial health. Low or insufficient income, as well as little or no education, may have led to insufficient nutrition and underinvestment in health (e.g. inadequate sanitary conditions, lack of medical advice and medicines, no recuperation time during illness). This may in turn have affected the general health level, and may have led to illnesses further weakening the general health level. Furthermore, for the working population there will be a dual causality between income and health. Bad health may lead to lower wage, lower-paid work, as well as spells of unemployment, which in turn may lead to low income and bad health.

That said, people who have a history of chronic obstructive pulmonary disease (COPD) or cardio-pulmonary problems are also thought to be particularly vulnerable on high air pollution days, and these types of health problems are more pronounced in high-income countries. The latter finding may have two reasons. First, these problems tend to occur later in an individual’s life, and since the population in high-income countries on average live longer than in low-income countries, they are more common in the former. Second, these diseases are to a certain extent life style related, i.e. they may be due to a mixture of excessive or unhealthy food intake, high stress levels, and lack of physical activity.
These insights would tentatively imply that when comparing income-groups (i.e. cross-sectional analysis) within a country, the association between pollution exposure and health is greater for lower income groups due among others to lower baseline health levels. When comparing high and low-income countries, on the other hand, baseline health’s effect on the size of the dose-response coefficient is ambiguous because not only is the general health level lower in the low-income country, but so is the occurrence of COPD and cardio-pulmonary problems and the average age.

5.4.3 Averting Activity and Exposure

This section specifically analyses the relationship between ambient pollution concentration, exposure, and averting behaviour. Unless a change in pollution leads to an equal-sized change in exposure for all individuals (i.e. \( \frac{dE}{dP} = 1 \)), ambient concentration will be a poor proxy for exposure to air pollution. In particular, we want to discuss whether there are reasons to believe that the effect of ambient pollution \( P \) on exposure \( E \) varies with the level of income \( Y \), i.e. \( \frac{d^2E}{dP \ dY} \neq 0 \).

Exposure can be defined as the time-weighted concentration of a given pollutant present in an individual’s breathing zone (Lipfert and Wyzga (1997)). To date, most dose-response studies have used readings of ambient air pollution concentration from central monitoring stations as a measurement of individual exposure to air pollution (and should then more accurately be named “concentration-response” studies), since this is usually the only measurement available. This implies that all individuals residing in one specific measurement zone are recorded with the same amount of exposure. This is not unproblematic, however, since exposure to air pollution may differ between individuals and between population groups.

An individual can choose whether to undertake measures to decrease the daily exposure to air pollution, so-called *averting* measures, and this decision will depend on constraints imposed by income, time, and information available, as well as on the level of pollution, the price of averting behaviour relative to other prices, the initial health level, and perceptions of risk and of own ability to heal. In some cases written or unwritten “rules” may impose additional constraints, as when certain population groups are prevented from living in certain areas of a city.

The concentration of pollutants in the air will typically vary at different locations in a city depending on the location of polluting industries, the geographic features of the
city in question, the wind directions etc. Individuals choose to live at different locations, as well as work in and travel through different areas and therefore be exposed to different amounts of pollution. In addition, a person’s choice of cooking and heating fuels used in the home, type of work, level of physical activity, as well as exposure while in traffic (e.g. choice of covering mouth and nose with cloth) will affect the dose of pollutants that he or she receives. Finally, certain types of work may imply particularly high exposure to local pollution at the work place. Little of this individual-specific information is usually available and used in dose-response studies and the number of outdoor monitoring sites providing estimates of exposure is often limited.

As discussed in section 5.3 exposure can be modelled as a function of both pollution level and averting behaviour, and averting activity as a function of pollution and income (variables not relevant for the argument have been left out). The generic form of the relationship can be expressed as follows:

\[ E = E(A(P, Y...), P) \]

The crossartial and partial derivatives of exposure with respect to income are given by the following expressions:

\[
\begin{align*}
\frac{d^2E}{dPdY} &= \frac{\partial E}{\partial A} \frac{\partial^2 A}{\partial P \partial Y} + \frac{\partial^2 E}{\partial P \partial A} \frac{\partial A}{\partial Y} + \frac{\partial^2 E}{\partial P \partial A} \\
\end{align*}
\]

\[
\frac{dE}{dY} = \frac{\partial E}{\partial A} \frac{\partial A}{\partial Y}
\]

We thus need to distinguish between income affecting behaviour and thus the association between ambient pollution on exposure (expression (a)), and income affecting behaviour and thereby the level of exposure when controlling for air pollution (expression (b)). In the former case income acts as a modifier variable, whereas in the latter it will act as a confounder if pollution is correlated with income.

Pollution and income are correlated if changes in pollution levels (either over time or cross-sectionally) are correlated with general income levels, and thus economic activity. There is a large literature on this issue under the broad topic of ‘the environmental Kuznets-curve’ (EKC). The central idea is that environmental quality deteriorates initially and then improves as an economy develops. Although the empirical support of this Kuznets-curve has been mainly found in cross-sectional studies, there has been a tendency to assume it to hold across time. Two major explanations given for the EKC are summarised by Dinda et al. (2000):

(i) use of environment as a major source of inputs and a pool for waste assimilation increases at the initial stage of economic growth, but as a country grows richer, structural changes take place which result in greater environment
protection; and (ii) viewed as a consumption good, the status of environmental quality changes from a luxury to a necessary good as an economy develops.\textsuperscript{64} Grossman and Krueger (1995), using cross-country city-level data, found support for the environmental Kuznets-curve association, with a peak for lighter air particles (smoke) corresponding to per capita GDP level of US$ 6151 (measured at 1985 US prices). Shafik (1994) estimated the turning point for suspended particulate matter to be at per capita GDP US$ 3280. Hence, it is quite plausible that pollution is correlated with income, but since the sign may vary according to the income level in question the bias that would result from omitting income may also vary in sign.

Turning to the effect modifier role of income, let us for the moment assume that increases in income over time, or cross-sectionally, lead to higher levels of averting activity (below we will return to reasons why this is a plausible assumption). A change in pollution will clearly tend to have a smaller effect on exposure for larger levels of averting activities, because it would otherwise not make sense to invest in averting activities. Furthermore, it seems reasonable to assume that a change in pollution will lead to a larger change in the amount of averting activities undertaken for higher income groups. Despite the possibility that the effect of averting activities on exposure decreases with higher levels of averting activities,\textsuperscript{65} the effect of pollution on exposure is therefore likely to decrease in income, i.e. expression (a) above takes a negative sign.

For a given level of measured ambient pollution, individuals with relatively low income may tend to live in more polluted areas because other things being equal pollution drives the housing prices down (Harrison and Rubinfeld (1978), Portney (1981), Crane et al. (1997)). Other factors are however usually not equal and, in particular, a central location in a city with short distance to various facilities tends to push up the house price. Nevertheless, empirical evidence suggests that within urban areas, the rich live in areas with less ambient concentration of air pollution (Maddison et al. (1995)). This finding suggests a possible negative sign of expression (b), although it is unlikely that the estimated effect of pollution on exposure will be biased due to omitting income from the analysis.\textsuperscript{66} Whether an increase in the level of air pollution implies a larger increase in

\textsuperscript{64} Dinda et al. (2000), p. 410.
\textsuperscript{65} This may be the case if individuals choose the most efficient activities first.
\textsuperscript{66} For a given level of measured pollution in a specific area, it is possible that a certain group of the population will move to less polluted neighbourhoods if their income increases. However, housing is likely to depend on relative rather than absolute income level, and income of a certain population group relative to the rest of the area's population is unlikely to be correlated with pollution over time. Hence, when looking specifically at choice of residential area, the estimated effect of pollution on exposure over time is
exposure for low-income groups due to the location of their houses, and hence a negative sign of expression (a), will in the short run depend on where the additional pollution originated. In the longer run the degree to which housing prices adjust to the new pollution levels, and the degree to which people move according to the changed prices will determine the exposure for different income groups. Since relative income is the most relevant variable when it comes to choice of residential area within e.g. a given city, overall income will not play an effect modifier role when comparing dose-response estimates across cities for this particular averting activity case.\(^{67}\)

The less expensive types of cooking and heating fuels used primarily in developing countries also tend to be the most polluting, which may imply a tendency for higher indoors exposure for lower income groups. However, within the lower income groups the level of use of these fuels may increase with increasing income up to the income level where other types of heating and cooking fuels may be afforded. An empirical study of total exposure (i.e. indoors and outdoors) to particulate matter (PM\(_{10}\)) in India (Saksena and Dayal (1997)) found that slum dwellers had the highest exposure both from indoors and outdoors air pollution. Indoor air pollution emission is only partly affected by an increase in outdoor air pollution levels, and could therefore be seen as an argument why different income groups may suffer different levels of exposure at a given level of air pollution – thus supporting the idea that expression (b) takes a negative sign. On the other hand, it is not unlikely that due to worse quality housing more outdoor air pollution finds its way into the houses of the low-income groups (Ostro (1994)). If an increase in the level of air pollution leads to a larger increase in outdoor exposure for low-income groups, however, then even if the same percentage of outdoor air pollution leaks into the houses of different income groups, the implication will be that increases in the level of outdoor air pollution lead to a larger increase in total exposure for low-income groups, and hence a negative sign of expression (a).

Some evidence also suggests that the work place of low-income individuals is more polluted than that of the better off, especially in developing countries,\(^{68}\) although this evidence is scarce.\(^{69}\) On the other hand, compensating-wage studies have found that

\(^{67}\) We assume that perpetuating former choices, e.g. deciding to continue living in the same area, is also a form of averting activity.

\(^{68}\) Personal communication with NILU (the Norwegian Institute for Air Research), survey from Cairo, Egypt.

\(^{69}\) It is possible, however, that the lower-income workers are employed in older types of industries where the pollutant types are better known and measured, whereas the exposure to pollutants of higher-income
workers are compensated by a higher wage when being exposed to larger health-risks. These latter studies, however, looked only at accidents and not morbidity. If there is a connection between higher air pollution levels locally at the workplace and an increase in overall air pollution levels, and if the workplace of the lower income groups were to be more polluted than that of higher income groups, then increases in the ambient levels of air pollution may imply a higher increase in exposure for the lower income groups and thus a negative sign for the (a) expression. Otherwise, this is an argument for why at a certain level of pollution, some income groups will suffer more exposure and therefore another reason for a negative sign for expression (b).

Finally, the lowest income groups typically have little education and may therefore also have less information about the damaging effects of pollution and the ways in which to decrease daily exposure. Schooling has been found to have an effect on the efficiency of producing health (Berger and Leigh (1989)), and is also correlated with income. A low level of education may have an impact on the amount of averting expenditure made and/or the efficiency of that expenditure, and therefore both on expressions (a) and (b). However, the direction of the effect is not clear. Less information and more uncertainty could in principle just as well lead to overcompensation of averting activities as well as underutilisation.

When comparing low and high-income countries, rather than low and high-income groups within a country, there is an additional reason why the effect of pollution on exposure may differ between income groups. Ostro (1994) draws attention to the fact that many low-income countries have a warm climate and the residents in these climates are therefore likely to spend a greater portion of their time outdoors. ‘If this is true’ he argues ‘it is likely that a given level of air pollution, everything else constant, would generate much greater health effects in poorer, warmer countries.’ (Ostro (1994), p. 49). This finding suggests that both expressions (a) and (b) take a negative sign.

In summary, averting measures imply a cost to the individual and low-income individuals may not be able to afford undertaking them. Other things equal, potentially larger exposure at a given pollution level may cause a higher number of, and more severe incidence of, illness and mortality in lower income groups. It is also possible that increases in measured ambient pollution levels lead to larger increases in exposure levels for the lower income groups. This suggests that we could expect a larger dose-response

workers in newer industries is less well-documented. Hence, this observation may in the longer run be
estimate in low-income groups than in high-income groups because ambient pollution levels sampled at central monitoring stations are used to measure pollution exposure.

Several empirical studies have found a relationship between income and exposure to various air pollutants. In a study of 3 American cities (Washington DC, Kansas City, St. Louis) Freeman (1972) found that persons with low incomes resided in lower quality air (the study looked in particular at the level of particulates and sulphur dioxide). Asch and Seneca (1978) discovered in their study of 284 cities in the US that exposure to particulates was highest in larger, more densely populated cities with lower incomes and education levels, and with higher proportions of non-white residents. They also investigated the intra-city variation in air quality in Chicago, Cleveland and Nashville and found that poorer census tracts experience higher pollution levels. Finally, they attempted to analyse the distribution of air quality changes (between 1972 and 1974), and found that the change in particulate pollution appeared to have been distributed progressively (i.e. larger percentage reductions in low-income areas), whereas no consistent association was found for sulphur dioxide. There are two main problems with the studies just presented. First, the finding that exposure is higher for lower income groups is based solely on exposure measurements by the monitor nearest the place of residence. Work and school location and other individual characteristics are not taken into account and may skew results. Second, when Asch and Seneca analysed the distribution of air quality changes they assumed that the relative income positions of the census tracts did not change significantly between 1972 and 1974. Hence they did not control for behavioural responses to the changes, such as moving in or out of the area. Furthermore, behavioural responses are likely to occur over longer time periods, and a two-year study period is not sufficient. In a study of exposure to ozone and particulates in the South Coast Air Basin of California Brajer and Hall (1992) based their estimate of exposure not only on location or residence, but also on mobility within the basin and time spent in various indoors and outdoors activities. They find, among others, that low-income residents suffer from worse air quality than those with incomes in the top 40 percent.

misleading.
5.4.4 Individual Behaviour and Measured Health Effects

Let the measured individual adverse health effect, $\tilde{H}_i$, be a function of the actual adverse health effect, $H_i$, and the income level of an individual, $Y_i$:

$$\tilde{H}_i = h(H_i, Y_i)$$

This section discusses whether the income level of an individual, $Y_i$, affects the relationship between the actual adverse health effect, $H_i$, and the measured one; $d^2\tilde{H} / dHdY \neq 0$.

The dose-response literature usually distinguishes between mortality and morbidity effects of air pollution. Death is clearly the endpoint of concern in the former, whereas in the latter case the focus is on non-fatal illness or specific symptoms. Mortality and morbidity will, however, only be correctly measured if all incidences are reported and thus captured in the statistics. It is not unlikely that the probability of being recorded may vary with individual characteristics, and more specifically with income group.

As there is only one way to count a death-occurrence, there is not much room for measurement error in the mortality rate. This is probably one of the reasons why studies of mortality from air pollution are much more numerous than morbidity studies. Nevertheless, to what extent all deaths are being registered is still an issue. It is plausible that a number of deaths among the poor, especially in developing countries, are not captured by official statistics (Todaro (1993)). If this is the case, the mortality rate among the poor may be underestimated.

Morbidity is defined by the US Public Health Service as 'a departure from a state of physical or mental well-being, resulting from disease or injury, of which the affected individual is aware' (Cropper and Freeman (1991)). There are several ways of classifying morbidity measures, but for the present purposes 3 characteristics are particularly useful: i) occurrence, ii) duration, and iii) severity. The occurrence of a disease can be measured by counting the number of reported cases. This type of morbidity measure does not distinguish the degree of severity or the duration of the disease. The typical morbidity measures that, although imperfectly, also take duration

---

70 The requirement that the individual be ‘aware’ of the health effect experienced is, however, not unproblematic. First, this awareness may require a certain change in condition. If the health of an individual is constantly below a normal level of well-being because of e.g. pollution, he may not be aware of it or willing to pay in order to avoid it, but society may still be willing to spend some money to improve his physical condition. In addition, an individual may well be ‘aware’ of the wrong symptom (e.g. mistaking physical with psychological symptoms).
into account (but not severity) will distinguish between chronic and acute illnesses. Finally, morbidity can be classified according to impairment of activity (i.e. severity), as well as duration. The latter can be divided into several categories. During a Restricted Activity Day (RAD) an individual can undertake some but not all of his normal activities, whereas during Work Loss Days (WLD) the individual is not able to engage satisfactorily in his employment. Bed Disability Days are those days where a person is confined to bed, either at home or at a hospital. These latter morbidity measures give an idea of how severe the illness is, and by being measured in days can also give an idea of duration, however, they reflect behavioural responses to health changes rather than the health changes themselves.

The traditional statistical measurements used in connection with dose-response studies of morbidity from air pollution have been WLDs (e.g. Zuidema and Nentjes (1996), Ostro (1983), Hansen and Selte (2000)) and measurements based on statistics from hospitals or clinics. The latter category includes both emergency room visits (e.g. Samet et al. (1981), Schwartz et al. (1993), Rossi et al. (1993)) and hospital admissions (e.g. Pope (1991), Burnett et al. (1992), Schwartz et al. (1996b), Vigotti et al. (1996)), either due to all-cause morbidity or due to incidences of specific illnesses or symptoms.

If there is a tendency for low-income individuals to work although being ill in order to earn enough to sustain a minimum level of consumption, then the health effect of air pollution for these income groups would not be recorded or at least be underestimated using WLD as adverse effects measurement. Furthermore, hospitals or other medical clinics may not be available for the lowest income groups or if available they may often be of lower quality than those visited by the richer part of the population, and visiting them may imply direct costs that lower-income individuals are less able to afford (transportation, food, medical bills, bribes etc.). Hence, also the hospital and emergency room statistics may tend to underestimate the effect on low-income individuals. If working during the illness period or failing to visit a hospital or emergency room increases the severity and extends the duration of the illness, then these low-income individuals may in fact experience a larger adverse health effect than the higher income groups.

\footnote{It is also worth noting that low-income families may not be able to afford to keep a parent at home to look after a sick child, or to pay someone to do so. Hence, studies that look specifically at morbidity in children using absenteeism from school or from day-care centres to pick up the adverse effect may also yield biased results, as children of poor families may not be absent although they are ill (e.g. Pönkä (1990)).}
5.5 Why Willingness to Pay May Vary with Income

Based on the discussion and evidence presented in the previous section we propose that it is not unreasonable to argue that higher exposure, worse initial health, and less investment in health may together or separately lead to larger adverse health effects from air pollution for lower income groups. In this section we will discuss what determines the marginal willingness to pay (MWTP), and more specifically in what way the level of income affects MWTP.

The economic value placed on the change in health impact is usually based on either some measure of MWTP for reduced morbidity derived either from direct questioning (contingent valuation), labour market studies or from the averting behaviour method, or on the cost of illness approach. As argued earlier, MWTP is the preferred value because it takes into account the effect of adverse health events on human wellbeing.

An individual’s willingness to pay for a change in the probability of an adverse health effect occurring \( V_H \), is the income (more generally wealth) that must be taken away from her, \( \partial Y \), to compensate her for the reduction in probability while keeping utility \( v \) constant, whereas the willingness to pay for a change in the level of air pollution \( \partial P \), is the income that must be removed to compensate for the reduction in pollution while keeping utility constant.\(^72\)

\[
V_H = -\frac{\partial v}{\partial Y} \frac{\partial H}{\partial Y} \quad (5.16a)
\]

\[
V_P = -\frac{\partial v}{\partial Y} \frac{\partial P}{\partial Y} = -\frac{\partial v}{\partial Y} \frac{\partial H / \partial P}{\partial Y} = V_H \frac{\partial H}{\partial P} \quad (5.16b)
\]

Here we see that the MWTP for reductions in pollution \( V_P \), is equal to the MWTP for reductions in adverse health effects \( V_H \) multiplied by the direct effect of pollution on health. Holding the numerators in the equations (5.16a) and (5.16b) constant, it is clear that if the marginal utility of income is decreasing in income then the denominator will be smaller for higher income groups and therefore the WTP higher. There is, however, no reason why the numerator should not change with income as well. In particular, if the marginal utility of health is decreasing with higher health levels, and if higher income groups tend to have higher health levels, then we would also expect the first term in the numerator of (5.16a), i.e. \( \partial v/\partial H \) to be smaller for higher income groups. On the other

\(^72\) The \( i \)-subscripts are suppressed since there is no room for misunderstanding.
hand, one can easily imagine that individuals with poor health, e.g. suffering from some chronic illness, are more used to pain and suffering and therefore experience less decrease in utility when they get an additional illness than do healthy individuals, i.e. marginal utility of health may be increasing in health. Furthermore, it is important to keep in mind that health is a multifaceted concept and although poor people may be worse off in some health regards (e.g. general levels of resistance and speed of recovery), in other health regards they may have an advantage (e.g. cardio-vascular problems related to obesity). Finally, as discussed in section 4.3, the second term in the numerator of expression (5.16b) may be decreasing in baseline health. It is thus so far not clear whether MWTP will increase in income. We therefore need to set out the full model in order to be able to say something more about the effect of income on WTP.

As before, we assume that the adverse health event $H$ is a function of exposure $E$, mitigating activities $M$ and baseline health $B$. Furthermore, exposure is a function of ambient air pollution $P$ and averting activities $A$. Finally, we assume that the individual derives utility from consumption $C$, and disutility from adverse health events $H$:

$$U = u(C, H(E(P, A), M, B))$$

(5.17)

Health is not something that can be bought directly, rather we spend money on averting ($A$) and mitigating ($M$) activities, and these in turn affect health. Income $Y$ is spent on consumption, averting activities, and mitigating activities (with prices $p_A$ and $p_M$ respectively):

$$C + p_A A + p_M M \leq Y$$

(5.18)

The individual chooses $A$, $M$, and $C$ so as to maximise utility subject to the budget constraint:

$$\text{Max} U[C, H(E(P, A), M, B)] \quad \text{s.t.} \quad C + p_A A + p_M M \leq Y$$

This yields the first-order conditions:

$$\frac{\partial U}{\partial H} \frac{\partial H}{\partial E} \frac{\partial E}{\partial A} = \lambda p_A$$

(5.19a)

$$\frac{\partial U}{\partial H} \frac{\partial H}{\partial M} = \lambda p_M$$

(5.19b)

The Lagrangian multiplier $\lambda$ can be interpreted as the marginal utility of income ($\partial U/\partial Y$), and the numerator of the MWTP expression (5.16b) can now be expressed in the following manner:
By rearranging expression (5.19a) and substituting it into expression (5.20), we can rewrite the MWTP for a reduction in pollution as follows:

\[
\frac{\partial v}{\partial P} = \frac{\partial U}{\partial H} \frac{\partial E}{\partial E} \frac{\partial P}{\partial P}
\]

By rearranging expression (5.19a) and substituting it into expression (5.20), we can rewrite the MWTP for a reduction in pollution as follows:

\[
V_p = -P_A \frac{\partial E}{\partial P} \frac{\partial P}{\partial A}
\]

The MWTP for a reduction in pollution is hence equal to the price of averting activity (which is determined by the market and can hence be assumed to be given and equal for all income groups) multiplied by the marginal rate of technical substitution of pollution for averting activity (\(\text{MRTS}_{pa}\)) in the production of exposure. Whether the MWTP differs between income groups therefore depends on whether the MRTS varies at different income levels.

There exist empirical studies of the relationship between indoor and outdoor particulate matter concentrations and the implications for personal exposure (e.g. Smith (1993), Tsai et al. (2000)). These studies, however, do not estimate the pollution-exposure function (i.e. the effect of ambient pollution level and of averting activities on the amount of exposure experienced by an individual at different levels of averting activity and of pollution), but rather the pollution-exposure correlation.

It seems reasonable to assume that for a given level of averting activity, exposure to air pollution is a linear function of ambient concentrations in the air. For averting expenditures, however, it is likely that the most efficient expenditures are undertaken first and that the more averting activities are undertaken the more costly or inefficient they become. On the basis of this reasoning we suggest that for a given level of pollution, exposure is decreasing in averting activities at a decreasing rate. Combining this information we can now draw the isoquants for different exposure levels as in figure 5.2 and 5.3. Clearly, higher-indexed isoquants imply higher levels of exposure.

If the efficiency of averting activities only depends on the pool of averting activities available, then the slope of the isoquants will be the same at all levels of pollution for a given level of averting activities, as in figure 5.2. If, on the other hand, the efficiency of averting activities in decreasing exposure also depends on the level of ambient pollution, the slope of the isoquant will no longer depend only on the level of averting activities. In particular, it seems highly plausible that at higher levels of pollution and starting from the same amount of averting activities, less increases in
averting activities will be required to hold exposure constant when pollution increases by a certain amount (figure 5.3). \(^7\)

Figure 5.2: Exposure isoquants when efficiency of averting activities independent of the pollution level.

Figure 5.3: Exposure isoquants when efficiency of averting activities depends on the pollution level.

Source: Own illustrations.

\(^7\) To illustrate this, we introduce a simple exposure equation and use time spent indoors (IT) as an example of averting activity. The units of exposure \(E\) are pollution concentration \(P\) multiplied by time \(T\) and can be represented by the following equation:

\[
\Delta E = P \times \Delta T
\]

Assuming that indoors exposure is half of the exposure during time spent outdoors (OT), we can rewrite the equation in the following manner:

\[
\Delta E = P \times OT + IT \times P / 2
\]

At an average pollution concentration level of 10 \(\mu g/m^3\) and with no averting activity exposure during a 10 hour period is 100\(\mu g/m^3\). In order to keep exposure constant when pollution concentration levels increase by 10 \(\mu g/m^3\), an individual will have to spend all 10 hours indoors. Had the pollution level risen by 10 \(\mu g/m^3\) from an initial level of 40\(\mu g/m^3\), the individual would only have had to spend 4 hours indoors. Hence, in this example less averting activity is needed at higher levels of pollution to counterbalance a given increase in pollution concentration levels. Figure 5.3 illustrates isoquants for exposure where the efficiency of averting activities depends both on the amount of averting activities undertaken (decreasing efficiency with higher amounts of averting activities), as well as on the level of ambient pollution (increasing efficiency with higher amounts of pollution). We observe that when moving from low-exposure isoquants to higher ones (i.e. from \(E_1\) to \(E_3\)) for a given level of pollution, \(\overline{P}\), the difference in slopes is even larger in figure 5.3 than in figure 5.2.
The implication of this discussion is that the MRTS is quite possibly higher at higher levels of averting activity. This, in turn, suggests that the effect of income on the MWTP for a reduction in pollution is of the same sign as the effect of income on the amount of averting activities undertaken. Assuming health is a normal good, and averting activities are what we buy to obtain better health, averting activities will tend to increase with higher income levels, implying that MWTP increases with income. Dividing this MWTP by the direct effect of pollution onto adverse health, which as argued in section 4 possibly is higher for lower income groups, means that the MWTP for a reduction in morbidity is also increasing in income.

Empirically, WTP has been found to be higher for individuals with higher income (e.g. Krupnick et al. (2000), Alberini et al. (1997)). The fact that WTP is increasing in income would be reflected in a positive income elasticity of WTP. There are still very few empirical studies of the income elasticity of WTP for reducing the risk of adverse health effects. Alberini et al. (1997) and Jones-Lee et al. (1985) estimated an income elasticity of about 0.4, the Loehman et al. (1994) study valuing acute health effects estimated it to be between 0.26 and 0.6, from the work of Biddle and Zarkin (1988) an income elasticity of WTP of 0.7 can be inferred, and Viscusi and Evans (1990) suggest a much higher income elasticity of 1.1. Empirical evidence of a negative elasticity has not been presented. Turning briefly to an area where related evidence can be found, namely the WTP for environmental improvements (where the health aspect of the improvement often plays a significant role), the income elasticity has for a number of data-sets been found to be positive but less than one (Kristrom and Riera (1996), Pearce (1980)). It is theoretically possible for the income elasticity of WTP for improved health to be less than 1 and still have an income elasticity of demand larger than 1, implying health to be a ‘luxury good’. The fact that most studies have found the income elasticity of WTP to be significantly less than 1, however, makes it difficult not to conclude that health is not a luxury good.

As mentioned briefly in section 5.2, the use of the WTP-method as a welfare measure has been criticised, mainly because it implies that an additional pound in the pocket of a rich person causes the same welfare gain as in the pocket of a poor person. This becomes problematic if we for example look at potentially Pareto-improving

\[ \sigma = \eta / \varphi \]

Clearly \( \sigma < 1 \) is compatible with \( \eta > 1 \). However, for \( \eta > 1 \) to still hold we see that the smaller the income elasticity of WTP the larger the price elasticity of demand, in particular \( \varphi \) must be larger than \( \eta \).
reallocation of air pollution. The rich can in principle compensate the poor for accepting
a shift in polluting industries or traffic from the areas where they live into the areas where
the poor live. Such a potential Pareto-improving shift would lead to an increase in
economic efficiency. The question remains, however, whether this shift is a better
allocation of resources. If one cares about the distribution of wealth among people and if
the potential compensation is never carried out in practice, then the answer may well be
'no'. Let us weigh the price the different income groups are willing to pay for a decrease
in the probability of falling ill (given in equation (5.16a)) with the marginal utility of that
money:

\[
V_H \frac{\partial v}{\partial Y} = - \frac{\partial v/\partial H}{\partial v/\partial Y} \frac{\partial v}{\partial Y} = - \frac{\partial v}{\partial H}
\]

This procedure gives us the marginal utility gain from a decrease in the
probability of falling ill, and by multiplying this with the effect of a change in pollution
level on the probability of illness we obtain the utility gains from a certain decrease in air
pollution levels. If net utility gain is used in order to make the allocation decision the
outcome may be rather different. Earlier in this section we argued that higher income
groups tend to have higher health levels. If the marginal utility of health is decreasing
with higher health levels, it then follows that the marginal utility gain from a decrease in
the probability of falling ill will be smaller for higher income groups.

We have thus seen that the marginal willingness-to-pay for a decrease in the
probability of adverse health effects will tend to be lower for lower income groups,
whereas the marginal utility gain may well be higher for this group. Which of these
should be used when making policy decisions will depend on what kind of policy is
considered, whether compensation may actually take place, and whether one cares about
the distribution of wealth among individuals.

5.6 Four Income-Related Challenges and Suggested Corrections

The level and distribution of income may pose problems to the use of dose-
response functions. In this section we will discuss four income-related challenges to the
valuation and cost calculation of health impacts:

i) omitted variable bias

ii) transfer challenge

iii) averaging challenge

iv) information bias
To each of these shortcomings we will suggest corrections or improvements on the existing method.

The usual reason for estimating the slope of dose-response functions in a population sample is to project the health effect of changes in ambient air pollution levels in the same or other population samples. If mitigating activity is an important explanatory variable for the effect of exposure to air pollution on health and is omitted from the estimated model, then this may lead to a biased coefficient estimate. We show that regressing the adverse health variable onto air pollution exposure and a constant only will give us a downwardly biased estimator of the direct effect of exposure on health, and an upwardly biased estimator of the total effect. Including income as an independent variable in the regression equation will give us an unbiased estimator of the total effect of exposure on health, whereas inclusion of mitigating activity yields an unbiased estimator of the direct effect.

Even if the coefficients of interest are correctly estimated, it may not be correct to transfer them to other situations or populations. Individual characteristics may affect the size of the estimated associations, i.e. they may act as effect modifiers. We suggest that one way of dealing with this challenge is to introduce moderator variables, among them income, in order to tailor-make the dose-response coefficient to the country in question based on local characteristics. The coefficients used for these moderator variables can be obtained by regressing estimated dose-response coefficients (from meta-analysis) onto these variables.

However, even correctly estimated coefficients for the relevant population will yield biased estimates of health costs when combined with the estimated WTP to reduce the risk of adverse health effects for the same population, if both the WTP and the relationship between exposure and health response are affected by the level of income. The reason for the bias is that the covariance between WTP and the effect of exposure on health at the individual level is ignored. With the income-sensitive slope estimate derived above and information about the income distribution in a country we show how to improve the cost estimates of mortality from air pollution by letting the WTP estimates also be income-sensitive.

Finally, errors in the measurement of exposure and of the adverse health effects can cause so-called information bias. We argue that the traditional health measurements may under-represent the health effect on the lower income groups, and that this may
imply an overall downward bias of the health effect of pollution due to measurement error. On the other hand, we propose that at a given level of measured air pollution there is a tendency for lower income groups to be submitted to higher amounts of exposure than the higher income groups. The exposure of higher income groups is therefore exaggerated as compared to lower income groups, and the estimated dose-response coefficient is exaggerated in the low income groups as compared to the higher income groups. Alternative measurements of adverse health effects and exposure are briefly discussed.

The four income-challenges to the estimation process are illustrated in figure 5.4.

Figure 5.4: Various ways in which income level can affect the estimation of health costs from air pollution.

Source: Own illustration.

5.6.1 Omitted Variable Bias

The slope of the dose-response function describes the change in human health we would expect from a unit change in pollution. We showed in section 5.3 that this dose-response coefficient is composed of the direct effect of exposure on health and the indirect effect via behavioural choice (mitigating activities). In this section we will demonstrate that by ignoring the role of income, the coefficient obtained by regressing adverse health incidences onto the level of air pollution exposure gives a downwardly biased estimate of the direct effect of exposure on illness, and an upwardly biased
estimate of the total effect. Income and mitigation expenditure data can help us obtain unbiased estimates of both these effects.

Let us simplify the model from section 5.3 by leaving out the baseline health variable.\(^75\) Suppose we have cross-section data on a given population. Adverse health event \(H\) (morbidity or mortality) is modelled as a function of exposure, \(E\), and mitigating activities, \(M\), and mitigating activity is a function of exposure and income, \(Y\).\(^76\) Notice that averting activities affect the amount of exposure experienced for a given level of ambient pollution, and since measured pollution levels are used as a measurement of exposure, the effect of income on averting activities is therefore discussed under the section ‘information bias’. With a sample of \(n\) observations, the regression equation on adverse health events can be expressed as:

\[
H_i = \beta_0 + \beta_1 E_i - \beta_2 M_i + u_i \quad i = 1, 2, \ldots n \tag{5.22}
\]

where \(u\) is a random error term and the \(\beta\)s are the coefficients to be estimated. The regression equation for mitigating activity can be expressed as follows:

\[
M_i = \alpha_0 + \alpha_1 Y_i + \alpha_2 E_i + \lambda_i \quad \text{assume} \quad M_Y > 0, M_E > 0 \tag{5.23}
\]

where \(\lambda\) is the random error and the coefficients are denoted with \(\alpha\). In the discussion below we have dropped the subscripts to leave the equations uncluttered, since there is no ambiguity about the range of the summations.

Substituting for \(M\) in equation (5.22), we obtain the reduced form of the model:

\[
H = \gamma_0 + \gamma_1 E - \gamma_2 Y + v
\]

where

\[
\begin{align*}
\gamma_0 &= \beta_0 - \beta_2 \alpha_0 \\
\gamma_1 &= \beta_1 - \beta_2 \alpha_2 \\
\gamma_2 &= \beta_2 \alpha_1 \\
v &= u - \beta_2 \lambda
\end{align*}
\tag{5.24}
\]

If income is omitted from the regression equation (5.24), we can express the estimated model as follows:

\[
H = \gamma'_0 + \gamma'_1 E + v' \tag{5.24'}
\]

The estimated slope of the dose-response function for the population studied using ordinary least squares (OLS) is:\(^77\)

\[^75\text{It is worth mentioning, however, that if baseline health affects the adverse health effect experienced but is omitted from the regression analysis, the income variable may be proxying for baseline health.}\]

\[^76\text{See footnote 51.}\]

\[^77\text{The estimate of the relationship described in (24') is}\]

\[
\hat{H} = g'_0 + g'_1 E
\]

where \(\hat{H}\) is estimated morbidity for a given exposure. OLS estimation implies finding the value for the
\[ g'_1 = \frac{\sum \tilde{E}\tilde{H}}{\sum \tilde{E}^2} \]  

(5.25)

\( \tilde{E} \) and \( \tilde{H} \) denote deviations of the variables from their sample means, \( \bar{E} \) and \( \bar{H} \):

\[ \tilde{E} = E - \bar{E}, \quad \tilde{H} = H - \bar{H} \]

Substituting for \( \tilde{H} \) from equation (5.22) into (5.25) we have the following expectation (\( \epsilon \)) of the estimator:

\[ e(g'_1) = \beta_1 - \beta_2 e \left( \frac{\sum \tilde{E}\tilde{M}}{\sum \tilde{E}^2} \right) \]

(5.26)

As the income variable is omitted from the analysis, the estimated effect of exposure on the mitigation variable is given by:

\[ e \left( \frac{\sum \tilde{E}\tilde{M}}{\sum \tilde{E}^2} \right) = \alpha_2 + \alpha_1 e \left( \frac{\sum \tilde{E}\tilde{Y}}{\sum \tilde{E}^2} \right) \]

(5.27)

Substituting equation (5.27) into (5.26) yields

\[ e(g'_1) = \beta_1 - \beta_2 \alpha_2 - \beta_2 \alpha_1 e \left( \frac{\sum \tilde{E}\tilde{Y}}{\sum \tilde{E}^2} \right) \]

(5.28)

Assuming that exposure and income are negatively correlated (as argued in section 5.4.3), the term in parenthesis on the right hand side in equation (5.28) will be negative. This implies that the expected estimated effect of exposure on health, when the income variable is omitted from the estimation model, will be larger than the total (gross) effect of exposure on adverse health and smaller than the direct (net) effect of exposure, i.e. \( \beta_1 - \beta_2 \gamma_2 < e(g'_1) < \beta_1 \).

Omitting income as a variable in the regression equation therefore leads to a \textit{downwardly} biased estimate of the direct (net) effect of exposure on illness. The reason the bias is downward is that when exposure increases, so does mitigating activity which is assumed positively correlated with income, and mitigation counteracts the effect of exposure. If we control for income, the direct effect of exposure is larger. Omitting parameter \( \gamma_1 \) that minimises the squared residuals. The latter can be written as follows:

\[ \sum u^2 = \sum \left( H - g'_0 - g'_1 E \right)^2 \]

When this expression is minimised with respect to \( g_0 \) and \( g_1 \) we obtain the following conditions (the so-called normal equations):

\[ \sum H = ng'_0 + g'_1 \sum E \quad \sum EH = g'_0 \sum E + g'_1 \sum E^2 \]

By solving this system we obtain the expression given in (5.25).
income as a variable in the regression equation furthermore leads to an upwardly biased estimator of the total effect of pollution on adverse health if high exposure is associated with low income, and if low income is associated with lower levels of mitigating activities and thereby higher adverse health effects.

If we had only country-level information for the different variables, and were either carrying out a time-series analysis for one country over time or a cross-country analysis, the effect of omitting income would be slightly different. In the time-series case, we would still expect exposure to be positively correlated with mitigating activities, however, the association between income and pollution exposure is not necessarily negative. On the one hand higher income can allow for more averting expenditures, on the other we saw in section 5.4.3 that whether pollution and income are positively or negatively correlated may depend on the stage of development and income level at which the country currently finds itself (i.e. the EKC discussion). In the cross-country case, we would not only expect the association between income and pollution exposure to be ambiguous and depend on the income range of our country sample, but the association between mitigating expenditure and exposure is now also less clear-cut. In particular, the supply side when it comes to mitigating activities (such as the health services and medicines available, and the possibility of taking sick leave) may determine the variations between mitigation undertaken in different countries rather than the demand for these which is affected by exposure.

**Avoiding a bias**

The structural model consisting of equations (5.22) and (5.23) has two endogenous variables; $H$ and $M$, and two exogenous variables; $E$ and $Y$. We will briefly explain which coefficients we can estimate correctly if we have data on mitigating activities, income, or both.

Since the variable $M$ is expressed in terms of exogenous variables in equation (5.23) (i.e. it is already a reduced form), $M$ and the error-term $u$ in expression (5.22) are independent and the OLS estimates will be the best linear unbiased estimates. By including $M$ in the regression, the exposure coefficient will therefore be an unbiased estimator of the direct (net) effect of exposure on adverse health. However, the variables $M$ and $E$ are collinear which implies increased sampling variances of the estimates.
With income, in addition to exposure data, available we can obtain unbiased estimators for the total (gross) effect of exposure on adverse health and income with regression equation (5.22).

We can use the income and mitigation variables and indirect least-squares (ILS) estimation to yield consistent estimators for the net and gross effect of air pollution exposure on health, as well as for income and mitigating activities. The reduced-form equation (5.23) and equation (5.24) satisfy the assumptions under which OLS estimators are consistent so that \( g_1 \) and \( g_2 \) are unbiased estimators of \( \gamma_1 (\gamma_1 = \beta_1 - \beta_2 \alpha_2) \) and \( \gamma_2 \) (\( \gamma_2 = \beta_2 \alpha_1 \)), and

\[
\begin{align*}
    a_1 &= \frac{\sum \tilde{Y} Y}{\Sigma \tilde{Y}^2} \text{ is consistent estimator of } \alpha_1 \\
    a_2 &= \frac{\sum \tilde{M} \tilde{E}}{\Sigma \tilde{E}^2} \text{ is consistent estimator of } \alpha_2
\end{align*}
\]

This suggests taking the ratio of the first two estimators as an estimate of the coefficient on mitigating activity \( \beta_2 \):

\[
b_{2,ILS} = \frac{g_2}{a_1}
\]

Furthermore, we can obtain an estimate of the direct effect of exposure on adverse health, \( \beta_1 \):

\[
b_{1,ILS} = g_1 - a_2 b_{2,ILS}
\]

Hence, by including income in the regression equation we can obtain unbiased estimators of the total effect of air pollution exposure on adverse health events (as well as of income), whereas by including mitigating activity we obtain unbiased estimators of the direct effect (as well as of mitigating activity). With data for both income and mitigating expenditure, we can obtain unbiased estimators for the direct (net) and total (gross) effect of air pollution exposure on adverse health events, and for mitigating activities and income.\(^78\)

However, having obtained unbiased estimators for the total and direct effect of air pollution does not necessarily imply that it will be correct to transfer either to another population, or to the same population at another point in time.

\(^{78}\) It is worth mentioning that baseline health and the level of education, which also influence health in our full version of the model, will tend to be positively correlated with income. Hence, to correct fully for the biases discussed in the previous section we would ideally need data on these variables as well. If they are
5.6.2 Transfer Challenge

Until recently, the dose-response relationships based on data from developed countries (mostly the US, but also Europe, see CSERGE et al. (1999)) were directly transferred to other countries and applied to air pollutants concentration data in order to estimate the mortality and morbidity rate due to pollution in these countries (e.g. Ostro (1994), Margulis (1992), Aunan (1997), Lvovsky (1998)).

We have already seen that omitting income level as a variable in the regression equation may lead to a biased estimate of the exposure-response association. However, even if our estimates of the direct or total effect of exposure on adverse health were to be unbiased, both are given for a certain level of baseline health, income, and education and it may therefore not be correct to transfer either to another study sample. We will now discuss the implications of transferring both the net and the total effect coefficients from a high-income to a low-income country (see example 1 in appendix 5.B for a simple illustration of the problem), while taking into account that not only income, but also education and baseline health levels will tend to be lower in the latter, and that the three are highly interrelated.\textsuperscript{79} Below we will discuss the transfer of exposure-response coefficients, however, it is important to note that the transfer will usually be of dose-response coefficients. In the case of the latter coefficients there is the additional issue of how well changes in pollution levels proxy for changing exposure levels, and whether there is a difference between its performance as a proxy in developing and developed countries. This issue will be discussed further in sections 5.6.4 and 5.6.5.

We have earlier (section 5.4.2) argued that the relationship between exposure, \(E\), and health, \(H\), may be affected by the initial health level of the individual, \(B\). As for the association between mitigating activity, \(M\), and illness, it may be affected by the level of education, \(K\) (assumed to influence the efficiency of the mitigation expenditure), baseline health, as well as the level of mitigation and exposure. Finally, we suggested that the effect of exposure on mitigation depends on the levels of income, education, and baseline health:

\begin{equation}
\beta_i = \frac{\partial H}{\partial E} \overset{f(B)}{=}
\end{equation}

not available, the income variable is likely to proxy for them and therefore be biased.

\textsuperscript{79} In addition, some of the cities with the highest air pollution levels in the world are developing country cities. If there are non-linearities in the effect of exposure on morbidity or mortality, then this may be another factor that has to be taken into account when transferring coefficients from one population to another.
Total effect: \[ \gamma_1 = \frac{\partial H}{\partial E} = \frac{\partial H}{\partial M} \frac{\partial M}{\partial E} \] where
\[ \frac{\partial H}{\partial M} = g(B,K,M,E) \]
\[ \frac{\partial M}{\partial E} = h(B,K,Y) \]

If there is a tendency for lower baseline health levels to increase the direct effect of an increase in exposure on adverse health events (i.e. \( \beta_1 \beta_0 < 0 \)), then this will tend to increase \( \beta_1 \) in the developing country as compared to the high-income country for a certain age group. If age is not controlled, however, lower average age in developing countries implies that a smaller proportion of the population is particularly vulnerable on high exposure days, as discussed in section 5.4.2. Since a higher proportion of older people reflects an overall better health level of a certain population, failing to control age will lead to a downwardly biased estimator of the effect of higher baseline health in decreasing the effect of exposure on adverse health.

We now turn to the indirect effect of exposure on adverse health via mitigating activities. We will discuss the effect of exposure on the amount of mitigating activities undertaken, as well as the effect of mitigation on adverse health, and whether these effects may vary between poor and rich countries.

As argued in section 5.4.1, an increase in exposure is likely to lead to a larger increase in the amount of mitigating activities undertaken by higher income groups (i.e. \( \frac{\partial M}{\partial E} \beta_0 > 0 \)). As for education, and in particular awareness of the link between pollution and health, we would expect that an increase in exposure on the amount of mitigating activities undertaken will be lower in countries with low education levels (i.e. \( \frac{\partial M}{\partial E} \beta_0 > 0 \)). Alberini et al. (1997) found in their Taiwan study that expenditure on mitigating activities is increasing with education and income. Findings by Alberini et al. (1997) further suggest that expenditure on mitigating activities increases with lower baseline health if the baseline health factors considered are chronic and prior respiratory illnesses. In a study of older people’s WTP in Canada Krupnick et al. (2000) also found chronic conditions to have a positive effect on WTP, although the associated \( t \)-statistic was very low. These studies were, however, both carried out in high-income countries and looked at conditions that have been linked to air pollution. Looking at the baseline health level more generally, it is not unlikely that in countries with low levels of baseline health other health expenditures will be prioritised rather than those that can be attributed to air pollution effects. It is furthermore possible that being used to endure bad health makes people less susceptible to adverse changes in their health. For these reasons we
believe it is likely that moving from a country with high levels of baseline health to one with low levels, the effect of changes in exposure on mitigating activity will be lower in the latter (i.e. $\Delta M / \Delta E < 0$). The empirical evidence is not clear when it comes to the association between age and the effect of exposure changes on mitigation expenditure, and age is another factor influencing baseline health. On the one hand, older people are more susceptible to air pollution, but on the other they have shorter life expectancy and less people depend upon their well-being. Overall we conclude that it is likely that a change in exposure has a smaller effect on mitigating activity levels in developing countries.

Turning to the effect of a change in mitigating activities on the adverse health effects, it is plausible that with a lower education level the efficiency of each unit of mitigation expenditure will be lower (i.e. $\delta H / \delta MK > 0$). Furthermore, the health system in many low-income countries is not of high quality so that the efficiency of each unit of mitigation expenditure may be lower also for this reason. On the other hand, if the effect of mitigation is decreasing with higher levels of baseline health (i.e. $\delta H / \delta MAB < 0$), higher average age, and/or higher levels of mitigation then the lower baseline health, average age, and mitigation level in the developing country may imply a larger effect of mitigating expenditures on health ceteris paribus. Finally, it is not unlikely that the effect of increasing mitigation will lead to a larger decrease in the adverse health effect at higher levels of exposure, and as we have seen in section 5.4.3 exposure levels in low-income countries will tend to be higher. It is therefore not clear whether the effect of mitigating activities in decreasing the adverse health effect is lower in poor than in rich countries.

There is no clear-cut conclusion from the discussion above as to whether the partial and total effects of air pollution exposure on adverse health events would be higher in the developing country, however, several arguments seem to indicate that they may be.

If our empirical analysis regresses adverse health events on exposure and a constant only, the constant will pick up systematic (non-stochastic) deviations not explained by exposure. It is not unreasonable to assume that the intercept will be higher in a developing country (denoted by subscript B in figure 5.5 below) than in an industrialised country (subscript A), i.e. that there is a higher number of some adverse health incidence per head (e.g. mortality) irrespective of the level of air pollution. In
addition, as we have argued above, it may well be that the slope of the exposure-response function (the total derivative of morbidity/mortality with respect to exposure) is steeper in the developing country. In figure 5.5 we can observe the implication of transferring the estimated coefficient from country A to country B. If the level of exposure in country B is at $E_B$, and we want to calculate the change in adverse health events by reducing exposure to a level in line with the WHO guideline, $E_{WHO}$, we see that the real decrease is $h_a - h_c$, whereas the transferred coefficient would lead us to estimate $h_a - h_b$. Hence, the transferred coefficient can be viewed as giving a lower bound to the actual health effect experienced.

Figure 5.5: Comparison of actual and transferred dose-response slope in a developing country

In summary, lower-income countries may experience larger adverse health effects from a certain increase in pollution than a higher-income country. This implies that even when the dose-response coefficient has been correctly estimated for a rich country, transferring it to a developing country may induce a downward bias. How to deal with this transfer challenge will be discussed next.

Introducing Moderator Variables

In developing countries the data required to carry out an analysis of morbidity and mortality effects from air pollution are very often not available, or are of poor quality. While waiting for higher quality time-series data to be gathered in developing countries, a second-best approach to the problem exists and has been used in analyses by Levy et al. (2000), Maddison and Gaarder (2001), and in the final part (chapter 6) of this thesis. The method consists of explaining variations in the regression coefficient between studies from
different cities (meta-analysis) not only in terms of stochastic effects but also in terms of baseline factors or 'moderator variables' that influence the slope of the dose-response function. The coefficients derived from such a meta-analysis can then be used to transfer epidemiological results into contexts that differ in terms of these moderator variables.

As we saw in section 5.3, the exposure-response function can be considered a reduced form of the health production function and the slope of the former then consists of the direct effect of exposure on health, as well as the indirect effect of exposure on mitigating activities and thereby on health. We have previously argued that the effect of exposure on mitigating activities may depend on the income level \( Y \), the level of education \( K \), as well as the level of baseline health \( B \) of the individuals surveyed. Furthermore, the effect of mitigating expenditure on adverse health may also be influenced by the level of education, exposure, and baseline health. The direct effect of exposure on illness (or mortality) may furthermore depend on baseline health. Since the observations are cities or countries rather than individuals, and the moderator variables are based on country averages it may also be interesting to control for the distribution of the moderator variables within each country. A distribution variable will be sensitive to non-linearities in the effect of the variable itself on the estimated coefficients, as well as to cross-section non-linearities in the effect.

With a set of empirically estimated total or direct effects of health with respect to exposure \((g_1 \text{ and } b_1\), respectively), as well as some measures of baseline health, exposure level, income per capita, education, income distribution \(DY\), baseline health distribution \(DB\), distribution of education \(DK\), and of exposure \(DE\) we can now regress these coefficients on the various moderator variables:

\[
b_1 = \chi_0 + \chi_1 B + \chi_2 DB + u
\]

\[
g_1 = \theta_0 + \theta_1 Y + \theta_2 E + \theta_3 B + \theta_4 K + \theta_5 DY + \theta_6 DE + \theta_7 DB + \theta_8 DK + e
\]

Note that if ambient pollution levels are used as a measure of exposure, then we have to take into account that averting activities may affect the path from pollution to exposure. In particular, the effect of a change in pollution on the amount of averting activities undertaken may depend on the pollution level, income, baseline health, and knowledge. It is also worth mentioning that we may allow for non-linearities in the relationship between the direct effect of exposure on health by including exposure as moderator variable in equation (5.29).
Two procedures are thus suggested in order to be able to obtain the coefficients necessary to tailor make the dose-response coefficients to local conditions. The first is to include mitigating activity as explanatory variable when performing the regression on the health production function (equation (5.22)) and thereby obtain an estimate of the direct effect of exposure on illness, and then to regress the estimated coefficient onto some measurement of the level of baseline health, and possibly its distribution in the population. The second, and more relevant procedure given the existing empirical studies, is to regress estimated slopes of the dose-response function onto income, education, and exposure and the distribution of these in the population, in addition to the already mentioned baseline health variables.

We mentioned in section 5.6.1 that the estimated dose-response coefficients are not identical to the total effect of exposure on adverse health. This will be important to keep in mind when interpreting the coefficients obtained from regression model (5.30).

If the country to which we would like to transfer the dose-response function has values for the moderator variables that lie within the range of those covered by the meta-analysis, we can, on the basis of the estimated coefficients of the moderator variables and data for the moderator variables of the relevant country, now tailor make the dose-response coefficient.

5.6.3 Covariance Challenge

In the previous section we saw that population characteristics may have an impact on the overall effect of pollution on health. A coefficient that has been correctly estimated for one particular population and setting will therefore not be correct when transferred unaltered to another setting with differing relevant underlying characteristics. However, even though we possess the correctly estimated coefficient for a population, as well as a correct estimate of the willingness-to-pay (WTP) to avoid being ill for the population as a whole, the health cost calculation from air pollution is not without its hazards. Dealing with these expected values as if they were independent may lead to wrong results when calculating the value of health changes if both the WTP to reduce illness and the health effect of pollution differ between income groups (see appendix 5.B, example 2 for a simple numerical illustration of the issue).
We will thus demonstrate in this section how the standard cost-calculation is incorrect due to the problem of using averaged measurements if income affects WTP (as argued in section 5).

The WTP of an individual \( i \) to reduce illness is denoted \( V_{H_i} \) and \( \hat{V}_{H_i} \) is an estimate of the WTP. Let \( \beta_{ii} \) denote the direct effect of pollution exposure on individual health, and \( b_i \) be the estimated effect. Finally, \( \sigma_{V_{H_i} \beta_{ii}} \) is the covariance between \( V_{H_i} \) and \( \beta_{ii} \) and \( \sigma_{\hat{V}_{H_i} b_i} \) is the covariance of their estimates.

The traditional way of calculating health costs is to multiply the estimated effect of air pollution on health with the estimated WTP, the relevant population \( N_i \), and the change in pollution concentration \( \Delta P \):

\[
\text{Traditionally Calculated Health Cost} = \Delta P N_i \hat{V}_{H_i} = \Delta H \hat{V}_{H_i}
\]

The actual health costs from a change (e.g. an increase) in air pollution are equal to the sum of the individual health costs, which in turn are equal to the effect of the increase in air pollution concentration on individual health, multiplied by the increase in air pollution concentration and the price the individual is willing to pay to avoid the health change.

\[
\text{Actual Health Costs} = \sum_{i=1}^{N_i} (\Delta P V_{H_i} (\Delta H_i / \Delta P)) = \Delta P \sum_{i=1}^{N_i} (\beta_{ii} V_{H_i})
\]

Using the definition for the expectation of the product of two random variables,\(^8^0\) as well as the cost expression derived in (5.16b), the relationship between the expected actual health cost and the traditionally calculated health cost estimated using the direct effect coefficient is given by the following expression:

\[
\text{Expected Actual Health Cost} = \Delta P \mathbb{E} \left( \sum_{i=1}^{N_i} (\beta_{ii} V_{H_i}) \right) = \Delta P N_i \mathbb{E} (V_{H_i}) (\mathbb{E} (\beta_{ii}) + \text{Cov} (V_{H_i}, \beta_{ii}))
\]

\[
\text{Estimated Health Costs} = \Delta P N_i \hat{V}_{H_i} b_i + \sigma_{\hat{V}_{H_i} b_i} = \Delta P N_i \hat{V}_{H_i} b_i + \Delta P N_i \sigma_{\hat{V}_{H_i} b_i}
\]

\[
= \text{Traditionally Calculated Health Cost} + \Delta P N_i \sigma_{\hat{V}_{H_i} b_i}
\]

Only if the covariance between the estimated willingness-to-pay and the estimated direct effect of pollution on health is zero will the estimation of the expected actual health

\(^8^0\) Let \( X \) be a random variable with mean \( \mu_X \), and let \( Y \) be a random variable with mean \( \mu_Y \). The expectation of the product of these random variables is defined as:

\[
\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y) + \text{Cov}(X,Y) = \mu_X \mu_Y + \mathbb{E}[(X - \mu_X)(Y - \mu_Y)]
\]
costs be equal to the traditional way of calculating the health costs. If, as we suggested earlier, low-income individuals tend to have a lower WTP and a larger direct adverse health effect from a dose of pollution, then low values of WTP will tend to be associated with high values of dose-response coefficients which in turn will imply a negative value for the covariance. In other words, the usual way of estimating the health costs will tend to overvalue the actual costs.

We now turn to the relationship between the real health cost and the health cost estimated using the total effect coefficient $\gamma_t$. By rearranging the expression for the total derivative of the health production function with respect to pollution given in (5.10) we can express the direct effect of pollution on illness as follows:

$$\frac{\partial H}{\partial P} = \frac{dH}{dP} - \frac{\partial H}{\partial A} \frac{\partial A}{\partial P} - \frac{\partial H}{\partial M} \frac{\partial M}{\partial P}$$  \hspace{1cm} (5.31)

Substituting this expression into equation (5.19b) we obtain the following expression for the MWTP for a reduction in pollution:

$$V_P = V_H \frac{dH}{dP} - V_H \frac{dH}{\partial A} \frac{\partial A}{\partial P} - V_H \frac{dH}{\partial M} \frac{\partial M}{\partial P}$$  \hspace{1cm} (5.32)

We can further substitute from the first-order-conditions in (5.19) and this gives us:

$$V_P = V_H \frac{\partial H}{\partial P} = V_H \frac{dH}{dP} + P_A \frac{\partial A^*}{\partial P} + P_M \frac{\partial M^*}{\partial P}$$  \hspace{1cm} (5.33)

Here the stars (*) symbolise the utility-maximising amount of averting and mitigating activities. Let $\gamma_t$ denote the direct effect of pollution exposure on individual health, and $g_i$ is the estimated effect. Furthermore, $a_i$ and $m_i$ denote the direct effects of pollution on averting and mitigating activities, respectively, and $\hat{a}$ and $\hat{m}$ are the estimated effects. Finally, $\sigma_{\gamma_t, y_t}$ is the covariance between $V_H$ and $\gamma_t$ and $\sigma_{\gamma_t, y_t}$ is the estimated covariance.

As before, the traditional way of calculating health costs is to multiply the estimated effect of air pollution on health with the estimated WTP, the relevant population, and the change in pollution concentration. In this case:

$$Traditionally Calculated Health Cost = \Delta PN g_{\gamma_t} \hat{V}_H = \Delta H \hat{V}_H$$

Based on expression (5.33), the relationship between the real health cost and the health cost estimated using the traditional dose-response coefficient is given by the following expression:
Expected Actual Health Cost

\[ E \Delta P e = \left( \sum_{i=1}^{N} (V_{H_i, T_i} + P_{A} a_i + P_{M} m_i) \right) \]

\[ = \Delta PN \left[ e(V_{H_i})e(V_{T_i}) + \text{Cov}(V_{H_i}, V_{T_i}) + P_{A} e(a_i) + P_{M} e(m_i) \right] \]

Estimated Health Costs

\[ = \Delta PN \left[ \hat{V}_{H} g_{1} + \sigma_{\hat{V}_{H}} \hat{a} + P_{A} \hat{a} + P_{M} \hat{m} \right] \]

\[ = \text{Traditionally Calculated Health Cost} + \Delta PN \sigma_{\hat{V}_{H}} + \Delta PN \left[ P_{A} \hat{a} + P_{M} \hat{m} \right] \]

As above, the covariance will be negative if the low-income individuals tend to have a lower WTP and a larger overall adverse health effect from a dose of pollution. The covariance will however not be of the same size as under (5.32), because income’s effect on the overall effect of pollution on morbidity will not be the same as on the direct effect. In contrast to the result when using the direct effect, the estimated health costs will not be correct even when the covariance between the WTP and the dose-response estimate is 0. The latter is due to the fact that the estimated costs have suppressed the changed expenditure on mitigating and averting activities undertaken by the population. This latter oversight will therefore counteract the effect of the missing covariance, and it is hence in this case not clear whether the estimated costs are exaggerating or undervaluing the actual costs.

It is worth noting that if we are interested in the utility gains from a decrease in pollution and multiply the estimated effect of pollution on health with the estimated marginal utility gains from the reduction in probability of falling ill of the population, then the neglected covariance would probably have been of a positive sign, i.e. both the marginal utility gain from decreased probability of illness and the effect of a certain pollution dose on health will tend to increase with lower income.

In summary, as long as the changes in expenditure on mitigating and averting activities are taken into account when required, the traditional way of calculating the health costs from air pollution increases will exaggerate the actual costs, whereas the utility losses from the increase in pollution will be undervalued when using the population estimates.
Adjusting the Willingness to Pay

Not only may the number of deaths associated with air pollution differ due to different levels and distribution of income in different countries, but the WTP to avoid these health effects may also differ with income.

There are several approaches to transferring estimates of WTP computed mainly in the United States to other countries. The most common approach in the literature corrects only for income differentials between the countries. An elasticity of WTP with respect to income is then either assumed (usually it is assumed to be 1.0), or an estimate from the valuation literature is used. A more sophisticated approach involves transferring a WTP-function from one country to another, using variables such as income, age, and chronic illness as moderator variables. The coefficient of the estimated WTP-equation is used along with mean values for the exogenous variables taken from the target country in order to estimate WTP for that country (Alberini et al. (1997)). This latter approach is similar to the one proposed for transferring the slope of the dose response function in section 5.6.1.

Given that the existing WTP estimates are taken almost exclusively from the US there is a clear need to adjust the WTP for income effects before applying the results to developing countries. The empirical findings suggest that the income elasticity of the WTP for reduced risks to life or for improved health is less than 1, however, this does not necessarily imply that the estimated elasticity is the correct one to use when adjusting the WTP measured in the US for income level before transferring it to a low-income country. Other factors that affect WTP may also be correlated with income, such as education levels, and these are not taken into account in this simple approach to WTP transfer. Low-income countries also tend to have low education levels, and if the level of education has a positive effect on the WTP this may exaggerate the difference in WTP at different income levels and therefore also the income elasticity of WTP. Empirical findings in the study by Alberini et al. (1997) suggest that education has a positive effect on the WTP. They concluded that in the case of the US and Taiwan the assumption of an elasticity of 1, which should cause an underestimation of WTP in Taiwan, actually worked well compared to the more advanced approaches discussed for exactly this reason.

There are two points that are important to note about the WTP. Firstly, if we assume that human well-being depends on lifetime consumption, then WTP to reduce risk will fall with expected lifetime consumption (Freeman (1993)). On the other hand, it could be argued that the preference structure (in particular the preference over time) and
risk perceptions of individuals may change with age and may influence the WTP in the opposite direction. It is therefore not clear that WTP will fall with age. Secondly, it is still not clear from epidemiological studies to what extent a person’s death is advanced by air pollution. If those who die from air pollution might otherwise have enjoyed a normal life expectancy, or if the WTP is not affected by life expectancy then no adjustment to the WTP is required.

The approach used to date adjusts the WTP for the GNP per capita at purchasing power parity in the relevant country using an average of existing empirical evidence on the income elasticity of WTP (e.g. Maddison and Gaarder (2001), Alberini et al. (1997)). This is an important first step in the right direction. There is, however, a problem with using these cost calculation methods because they are using average WTP estimates. We have earlier in this section argued that if both the WTP to reduce illness and the health effect of pollution differ between income groups, then using these averages yields wrong results when calculating the cost of health changes even if both the estimated direct effect of pollution on health and the average WTP to avoid being ill were correctly estimated.

Based on the estimated value of statistical life (VOSL) for air pollution in one country (e.g. $ 2 million in the U.S, $800,000 to $2.5 million in Canada), the estimated income elasticity of WTP, and the average income in the different income groups, the average WTP for the different income groups can be calculated. These measurements can then be combined with the dose-response coefficients for each subgroup in order to calculate the costs from air pollution in income sub-groups of the population and add them up to obtain a more realistic cost measure.

5.6.4 Information Bias

In the previous section we saw that even though we have correctly estimated the mean direct effect of pollution on health, it may be an inappropriate level of aggregation when the WTP of different income groups differ. However, even if the sub-groups and

---

81 Empirical work is still limited in this field. Jones-Lee (1989) found that at least towards the end of the life cycle WTP is a decreasing function of age and therefore of remaining life expectancy. Subsequent work by Jones-Lee et al (1993) failed to support this initial finding, and Alberini et al. (1997) found that age has no substantial effect on WTP. Recent empirical findings in a study by Krupnick et al. (2000) found age to have only a modest negative impact on mean WTP. Mean WTP at age 75 was found to be approximately 75 percent of its peak value (reached at 54).

82 The value of a statistical life (VOSL) is the WTP for a small change in risk aggregated over a large number of individuals. This is not the same as valuing an actual life.


84 Krupnick et al. (2000).
their WTP are explicitly considered, bias due to errors in measuring the study variables (information bias) may still occur. Two sources of information bias will be discussed here; i) error in measuring the adverse effects, and ii) error in measuring exposure. In particular, the former error is due to adverse effects not being measured directly but rather through the behavioural (mitigating) response to it, which we have argued is affected by income. The latter is due to exposure not being measured directly but rather by proxy that does not take the behavioural (averting) response into account.

**Errors in Measuring the Adverse Effects**

We argued in section 5.4.4 that for the same amounts of illness (or even mortality) it is possible that the adverse health measurements traditionally used capture less of the illnesses experienced by the lower income groups (see appendix 5.B, example 3 for a simple numerical illustration).

Below we denote the measured individual adverse health effect as \( \tilde{H}_i \). If the only health effects captured are those that are costly to the individual, because of for example income loss due to work absence or hospital costs (direct costs and cost of lost earnings), or even a visit to a medical centre or burial costs, then the income level of an individual, \( Y_i \), may affect whether the actual adverse health effect, \( H_i \), he or she experiences is being measured.\(^{85}\)

\[
\tilde{H}_i = h(H_i, Y_i) \quad (5.34)
\]

Hence, in addition to potentially affecting the actual health effect, as we discussed in section 5.4.1, the income level may have an additional effect on the estimated health effect as seen in section 5.4.4.

Let us as an example look at work-loss-days (WLD) as a measure of morbidity, and assume that it is a function of the actual adverse health event, \( H_i \), and a measurement error, \( \phi_i \). The measurement error is in turn assumed to be a decreasing function of income, i.e. \( \phi(Y) \) with \( \phi < 0 \). The actual total health effect is modelled as a function of a constant, exposure to air pollution \( E_i \), income \( Y_i \), and an error-term, \( \nu_i \) (see the reduced form expression (5.28)). Let us, in order to simplify the argument, assume that we have a perfect instrument for income, \( Y^\dagger \), which is uncorrelated with the measurement error term.

---

\(^{85}\) Other individual characteristics, such as reluctance to admit sickness, fear of hospitals etc. may also affect the empirical measure. However, these characteristics are not important for the point we are making in this paragraph and are therefore left out of the discussion.
Combining these relationships then gives the following relation between the observed variables:

\[ \tilde{Y}_i = H_i + \phi_i = \gamma_0 + \gamma_1 E_i + \gamma_2 Y_i^l + \phi_i \quad (5.35) \]

where \( e(E, v_i) = 0, e(Y_i^l, v_i) = 0, e(Y_i^l, \phi_i) = 0, e(E, \phi_i) = 0 \)

Let us assume that there is no sick-pay so that only individuals who can afford it stay at home to recuperate, and that the less income an individual has, the larger the fraction of the sickness time he or she actually works. Let us define the existence income \( Y_{\text{min}} \) as the income level at which an individual can exactly survive given that he works the entire period (i.e. does not lose any work days), \( \tilde{Y} \) is the critical income level above which a person can afford to recuperate fully before returning to work, \( x_i \) is the amount of the working days during which the individual is ill that he works (i.e. \( 0 \leq x \leq 1 \)), and \( \varepsilon \) is a random error. We can then express the measurement error in the following way:

\[ \varepsilon_i = \varepsilon(Y_i^l, x_i) = 0 \quad (5.36) \]

The OLS estimator of \( \gamma_i \) in equation (5.35) is then for the two income groups given by:

\[ e(g_i) = \gamma_1 - x_i e\left( \frac{\sum E \tilde{H}}{\sum E^2} \right) = \gamma_1 (1 - x_i) \quad (5.37) \]

Hence, for any individual with an income level below the critical income the coefficient estimate of the dose-response function will be downwardly biased, whereas it will be correctly estimated for those with an income level above the critical level. If, due to their choice of working, the lower income groups extend the duration and/or severity of the illness, i.e. \( \gamma_i (x) \) and \( \gamma_{ix} > 0 \), then the actual adverse health event experienced by this group is larger than for the higher income groups, which would imply an even larger downward bias due to measurement error. For the relevant population as a whole the coefficient will be biased downward.\(^{86}\)

In many dose-response studies, e.g. the mortality studies, the researcher is interested in the effect of a change in air pollution on the percentage increase in daily events, rather than the absolute increase which we modelled above. If the fraction of the illness time spent working is not affected by exposure, then the fact that measured health

\(^{86}\) In reality, it will be difficult to say anything general about the sign of the bias. Even if lower income groups work during illness, it is possible that higher income groups report that they are absent from work.
is downwardly biased as compared to actual health, will not imply that the dose-response coefficient is downwardly biased:
\[
\frac{\partial \ln(H(1-x))}{\partial E} = \frac{\partial (\ln H + \ln(1-x))}{\partial E} = \frac{\partial \ln H}{\partial E}
\]
However, if we assume that the size of the fraction depends on the amount of sickness time experienced, in particular, that at a given income level the longer the duration of the sickness the larger the fraction of the sickness time worked (i.e. \( \phi_i = e_i - H_i x_i(H_i) \) with \( x_i(H) > 0 \)), then the dose-response coefficient will also be downwardly biased:
\[
\frac{\partial \ln(H(1-x(H)))}{\partial E} = \frac{\partial (\ln H + \ln(1-x(H)))}{\partial E} = \frac{\partial \ln H}{\partial E} + \frac{\partial (\ln(1-x(H)))}{\partial E} < \frac{\partial \ln H}{\partial E}
\]

**Alternative Ways of Measuring Adverse Effects**

In order to avoid the measurement challenges discussed in the previous section, we require measures of health impacts that are independent of behavioural choices, instead of hospital days, work loss days, and sick days, which are not.

There are an increasing number of studies that now use self-reported illness and health symptoms as a way to circumvent this problem (e.g. Propper and Upward (1992), Alberini et al. (1997), Ostro (1987)). This method has obvious drawbacks, as different individuals will have different levels of tolerance, and different ways of communicating an experience of bad health. Whether there may also be significant differences between income groups is not clear. Tentatively, a person who has grown up under poor conditions may have a higher level of tolerance than one that has not (Sen (1984)), and also less knowledge of illness symptoms.

An alternative way of capturing, although imperfectly, the adverse health effect from air pollution on those who continue working during their illness, is to measure their loss in productivity at work. The idea of a technically determined relationship between health and labour effort (or productivity) is not new. This hypothesis - often called the efficiency wage hypothesis (Stiglitz 1974) - implies, among others, that increasing poor workers wages will increase their productivity at work because they can afford better nutrition and improved health. There is much empirical evidence that health has an impact on productivity and earnings (e.g. Luft (1975), Deolalikar (1984), Rizzo et al. (1996), Bartel and Taubman (1979)). In a perfectly functioning market economy one would therefore look at the wage rate which would reflect the marginal productivity of due to illness even if this is not true.
the worker. However, by merely looking at the effect of illness on wages the direct effect of illness on the wellbeing of an individual is ignored. For many types of work, furthermore, the marginal productivity of a worker is not identifiable, and most types of wages are not flexible enough to give any measure of fluctuations in short-term productivity. Wages also may be as much institutionally determined as they are determined by individual productivity. Finally, many factors influence productivity other than health and many factors influence health other than air pollution. Hence, in addition to yielding at most an imperfect adverse health effect measurement this approach is highly time and information intensive.

Medical monitoring of the sample examined would clearly give the most objective and unbiased measure, but will for the purpose of most studies be far too expensive, as well as resource and time demanding. Clearly, there is the usual cost versus accuracy trade-off, but once a thorough study of the health impact of air pollution on different income groups has been carried out its results may be of great use to future benefit transfer studies.

**Errors in Measuring Exposure**

In section 5.3 we saw that exposure \((E)\) can be modelled as a function of pollution \((P)\) and averting activity \((A)\). However, we simplified matters somewhat by presenting the relationship in a linear form. It seems more realistic to model averting activities as affecting the effect of pollution on exposure, i.e. as affecting the size of the regression coefficient of pollution. Averting activities, in turn, is a function of among other factors income \((Y)\) and pollution:

\[
E_i = \psi_1(A_i)P + \xi_i \quad \text{with} \quad \partial \psi_1 / \partial A < 0 \quad \text{and} \quad e(P\xi_i) = 0 \quad (5.38)
\]

\[
A_i = \delta_1 Y + \delta_2 P + \eta_i \quad \text{assuming} \quad 0 < \delta_1, \delta_2 < 1 \quad \text{and} \quad e(P\eta_i) = 0, e(Y\eta_i) = 0 \quad (5.39)
\]

In section 5.6.1 we argued that adverse health events can be expressed as a function of exposure and mitigating activities:

\[
H_i = \beta_0 + \beta_1 E_i - \beta_2 M_i + u_i \quad \text{with} \quad i = 1,2,\ldots,n \quad (5.40)
\]

We also derived the following reduced form of the model:

\[
H_i = \gamma_0 + \gamma_1 E_i - \gamma_2 Y_i + v_i \quad \text{where} \quad \begin{align*}
\gamma_0 &= \beta_0 - \beta_2 \alpha_0 \\
\gamma_1 &= \beta_1 - \beta_2 \alpha_2 \\
\gamma_2 &= \beta_2 \alpha_1 \\
v &= u - \beta_2 \lambda
\end{align*} \quad (5.41)
\]
If we substitute expression (5.38) for exposure in equations (5.40) and (5.41) we see that the estimated effect of exposure using pollution as a proxy will be the following:

\[ e(h_t) = \beta_1 \psi_t \]  
\[ e(g_t) = \gamma_1 \psi_t \]  

(5.42a)  
(5.42b)

If we assume that a person who does not undertake any averting activities will be exposed to the measured amount of ambient air pollution and that the more averting activities undertaken the less air pollution the individual will be exposed to (i.e. that \( 0 \leq \psi_t \leq 1 \) and \( \frac{\partial \psi_t}{\partial A} < 0 \)), then for the population as a whole using pollution as a measure of exposure will lead to a downwardly biased estimate of the effect of exposure on adverse health. Furthermore, from equation (5.39) we see that the amount of averting activity is an increasing function of income, and hence the downward bias in the estimated effect of exposure on adverse health when using pollution to measure exposure will also be increasing in income.

In reality, it will be difficult to say anything general about the sign of the bias. Even if we could assume that lower income groups are exposed to higher amounts of air pollution than higher income groups, it could well be that the air pollution detector exaggerates exposure for all members of the relevant population. Alternatively, if it is located in a place of average exposure, it could underestimate exposure for some and exaggerate it for others.

When comparing the use of centrally measured air pollution as a measure for exposure in high and low-income countries, two aspects that may partly work in the opposite directions are worth considering. On the one hand, if exposure relative to measured pollution is less in developed countries than in developing countries, using pollution as a measure of exposure will exaggerate the exposure-response coefficient more in developing than in developed countries (alternatively, it will undervalue it less). On the other hand, in many developing countries indoor air pollution, which is another source of exposure, is a much higher health risk than outdoor air pollution. If it is vulnerable people (low baseline health levels, or of higher age) who tend to die or get ill from air pollution, and if indoor air pollution does not covary with outdoor air pollution from day to day, then the exposure-response association may be much larger but not be captured by a study using outdoor pollution as exposure measure. In other words, those who are vulnerable to outdoor air pollution may already have died or become ill from indoor air pollution. The exposure-response coefficient may in fact be larger in a
developing country than a developed one, whereas the concentration-response coefficient is lower.

**Alternative Ways of Measuring Exposure**

In section 5.4.3 we mentioned that dose-response studies have tended to use ambient air pollution concentration measurements as measure of individual exposure to air pollution, and that this may be problematic if exposure to air pollution of different individuals differs.

Exposure can be measured indirectly by combining micro-environmental monitoring or modeling with questionnaires or other information on human activities, or directly by biological markers and personal monitoring.

For simplicity we can divide the issue of indirect exposure measurement into 3 sub-issues: i) accuracy of the measurement in the micro-environment, ii) presence of the person in the micro-environment, and iii) individual characteristics (e.g. activity level, smoking habits).

Micro-environmental measurements involve monitoring air pollutant concentrations in the locations where exposure takes place. The ability of a given air monitoring station to represent population exposure estimates over a certain geographical area depends on the physical and chemical properties of the pollutants in question and their resulting spatial variability (Lipfert and Wyzga (1997)). It is, however, clear that greater spatial density of monitoring will improve the measurement of outdoor exposure. As for indoor and workplace exposure, the issue is rather one of measuring exposure in these micro-environments at all, rather than improving the accuracy. In situations where it is not feasible with direct measurements, micro-environmental concentrations are estimated using concentration models, which are based upon the physics and chemistry of the environment. The quality of the output from models depends on the assumptions, the degree to which factors are identified and specified, and the quality of the data.

The more information available concerning the amount of time an individual spends in each micro-environment, the more precise will be the exposure measurement. Furthermore, individual habits such as smoking, and individual level of activity will influence the amount of individual exposure to certain air pollutants. The more relevant information about the individual we have, the more precise will therefore be the exposure measurement. This type of information is usually obtained by questionnaires. Data based
on questionnaires are always vulnerable to the quality of the questionnaire design, as well as to a person’s ability to recollect.

Personal monitoring provides direct measurements of the concentration of air pollutants in an individual’s breathing zone, whereas biological markers refer to cellular, biochemical, or molecular measurements obtained from biological media such as human tissues, cells, or fluids and are indicative of exposure to environmental chemicals. (NRCB (1991)).

There are two main drawbacks with the direct method. First, in large field studies it has proven difficult to get sufficient numbers of people that are willing to carry the samplers or to obtain physiological samples from subjects. In addition, both procedures are time consuming and expensive. Second, these direct monitoring methods do not supply information on the physical environment (e.g. temperature, humidity, ventilation) in which exposures occur or on the factors (e.g. emission rates) controlling concentrations in those environments. In addition, the relationships among biological marker concentrations and air concentration of a pollutant are poorly understood. This type of information is necessary if effective averting and mitigating measures are to be found, and if efficient pollution regulation policies are to be designed.

5.6.5 Overall Measured Income Effect

We have seen that the way in which both exposure and health events are measured in various studies may bias the estimated dose-response coefficients. In particular, we have seen that if the only health effects captured are those that are costly to the individual, the health effect on low-income individuals may be underestimated and therefore the coefficient for the population as a whole will be downwardly biased. If measured ambient pollution is the correct measure of exposure only for low-income groups who cannot afford much averting activities, whereas it exaggerates exposure for high-income groups then there is a downward bias in the estimated dose-response relationship for the latter group. Here again, the coefficient for the population as a whole will be downwardly biased.

Let us first investigate the measured effect of exposure on health and how it varies with the income level. If both the exposure and health effects are measured with error, then by substituting expressions (5.36) and (5.38) into (5.35) we obtain the following formula:
\[ \hat{H}_1 = \mu_0 + \mu_1 P + \mu_2 Y_x^I + \kappa_i \]

and \( e(P\kappa_i) = 0, e(Y^I\kappa_i) = 0 \) \hspace{1cm} (5.45)

where

\[ \begin{align*}
\mu_0 &= \gamma_0 (1 - x_i) \\
\mu^E_1 &= \gamma^E_1 \psi_1 (1 - x_i) = (\beta_1 - \beta_2 \alpha_2) \psi_1 (1 - x_i) \\
\mu_2 &= \gamma_2 (1 - x_i) \\
\kappa_i &= (\gamma_1 \xi_i + \nu_i) (1 - x_i) + \epsilon_i
\end{align*} \]

The OLS estimator of \( \mu^E_1 \) in equation (5.45) is then given by:

\[ e(m^E_1) = \gamma^E_1 \psi_1 (1 - x_i) \]

Differentiating this estimator with respect to income we obtain:

\[ \frac{d(e(m^E_1))}{dY} = \frac{\partial \gamma^E_1}{\partial Y} \psi_1 (1 - x) - \frac{\partial x}{\partial Y} \gamma^E_1 \psi_1 + \frac{\partial \psi_1}{\partial Y} \gamma^E_1 (1 - x) \] \hspace{1cm} (5.46)

Since the fraction of the working days during which the individual works while being ill is decreasing in income (i.e. \( x_y < 0 \)), and the fraction of air pollution that is averted is increasing in income (i.e. \( k_y > 0 \)), it is not clear what the net effect of the measurement errors is. What is clear, however, is that unless the two measurement errors cancel each other out the estimated effect using the traditional measurements will be biased. We have further seen that the total effect of actual exposure on actual adverse health will not unlikely be decreasing in income because higher income groups have higher baseline health and are able to afford more mitigating expenditures. To summarise, it is possible that the total effect of exposure on adverse health is larger in lower-income groups, however, the measured effect could be either lower or higher. When comparing high and low income countries, the actual exposure to air pollution may be partly uncorrelated with ambient concentrations (due to indoor pollution), something which could imply that both the measurement used for adverse health and exposure may induce a downward bias in the estimated effect.

Let us move on to the effect of pollution on health, which is what is usually measured, and discuss how it varies with the income level. The problem of measuring exposure as opposed to ambient concentrations is thus not an issue in this context, however, the health effect is still assumed to be measured with error. As presented in equation (5.15) section 5.3, pollution affects both mitigating and averting activities, and also has a direct effect on health. We denote the estimator of the total effect of pollution on adverse health \( \gamma^E_1 \) in order to distinguish it from the estimator of the total effect of
exposure on adverse health, $\gamma^E$, from the previous page. The OLS estimator of $\mu^P$ is then given by:

$$e(m^P) = \gamma^P (1 - x_i)$$

Differentiating this estimator with respect to income we obtain:

$$\frac{d(e(m^P))}{dY} = \frac{\partial \gamma^P}{\partial Y} (1 - x) - \gamma^P \frac{\partial x}{\partial Y}$$

$x$ is decreasing in income and the measurement error will therefore lead to a downward bias in the lower income groups. The downward bias will tend to be even larger when comparing low-income and high-income countries because in the former the social welfare system is not extensive (see section 5.4.4 for a further discussion on the issue). On the other hand, the effect of pollution on health may be decreasing in income, not only due to the fact that exposure may have a larger effect on adverse health in lower income groups, but in addition because exposure may be higher in lower income groups for a given level of pollution. Once more, however, when comparing developing and developed countries indoor air pollution may blur the picture. It is therefore not clear whether the net measured effect will be higher or lower in lower income groups.

In table 5.1 below an attempt has been made to summarise the main biases discussed in this chapter. For the benefit of the overview, we have made the following simplifying assumptions. First, we are assuming that only low-income people in poor countries are exposed to the equivalent of outdoor-monitoring measures of air pollution. There will therefore be less upward bias the poorer the income group or country under consideration. The facts that indoor air pollution concentration in some instances can be higher than outdoor air pollution, and that some areas may have higher levels of outdoor concentration than what is measured by the central monitoring station are being ignored. Furthermore, the possibility that actual exposure may be partly uncorrelated with ambient concentrations (due to indoor air pollution) is assumed not to have major implications. Finally, we are assuming only high-income groups/countries can afford to be fully measured by the health measure. The downward biases will hence be larger the poorer the income group or country under consideration.
Table 5.1: Summary of Estimation Biases related to Income.
(Note: a parenthesis indicates that the bias is not believed to be of major significance, whereas an iteration of a sign suggests that the bias is believed to be highly significant.)

<table>
<thead>
<tr>
<th>Expected Signs of Biases When Failing to Take the Effect of Income into Account</th>
<th>In-Country Bias, Rich Country</th>
<th>In-Country Bias, Poor Country</th>
<th>Between-Country Transfer Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Pollution Concentration as Measurement for Actual Exposure</td>
<td>+ + (information bias)</td>
<td>+ (information bias)</td>
<td></td>
</tr>
<tr>
<td>2. Effect of Actual Exposure on Actual Direct Health Effect</td>
<td>- (omitted variable bias)</td>
<td>- (omitted variable bias)</td>
<td></td>
</tr>
<tr>
<td>3. Effect of Actual Exposure on Actual Total Health Effect</td>
<td>(+) (omitted variable bias)</td>
<td>(+) (omitted variable bias)</td>
<td></td>
</tr>
<tr>
<td>4. Relationship Between Measured and Actual Health Effect</td>
<td>- (information bias)</td>
<td>- (information bias)</td>
<td></td>
</tr>
<tr>
<td>Pollution-Response Relationship (i.e. 2 or 3, and 4 (1 is irrelevant))</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Exposure-Response Relationship (i.e. 1, 2 or 3, and 4)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5. Willingness-To-Pay multiplied by Direct Dose-Response Coefficient</td>
<td>+ (covariance bias)</td>
<td>+ (covariance bias)</td>
<td></td>
</tr>
<tr>
<td>6. Willingness-To-Pay multiplied by Total Dose-Response Coefficient</td>
<td>? (covariance bias and suppression of averting and mitigating activities costs)</td>
<td>? (covariance bias and suppression of averting and mitigating activities costs)</td>
<td></td>
</tr>
<tr>
<td>Costing of the Estimated Dose-Response Relationship</td>
<td>?</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

The important implication from the present discussion is that although lower income groups are likely to experience a larger adverse health effect from a certain change in measured ambient pollution levels, this will not necessarily show up in an empirical study. This, in turn, suggests that carrying out studies in situ in the case of developing countries does not necessarily lead to more correct estimation of the health costs from air pollution than does transferring dose-response coefficients from high-income countries.

5.7 Summary and Conclusion

This paper has discussed a number of reasons why ignoring the role of income may mislead the quantification and valuation process of the health effects from changes in ambient air pollution and suggested a few improvements.

---

87 We are assuming the transfer is made from a high to a low-income/developing country.
We started out by introducing a distinction between the *probability* of illness occurring and the *actual* illness event. We have shown that different types of health events may to a varying extent allow for different individual responses, with the implication that the measured coefficient of the dose-response function may vary according to how the health event is defined. The main implication is that it will be misleading to look only at the probability of falling ill if pollution also affects the path/duration of the sickness and if we are interested in the full health effect of air pollution. When discussing the effect of income on the valuation process it will hence be important to keep in mind that this depends on the adverse health event we choose to analyse.

After this clarification followed a discussion, in part supported by empirical evidence, suggesting that low-income groups may experience larger adverse health effects from air pollution exposure than high-income groups. We argued that income can affect the extent and type of health impact through its influence on the amount of mitigating behaviour undertaken, as well as on the baseline health of an individual, although the direction of influence is not clear-cut. In addition, we suggested that the usual way of measuring illness and exposure to air pollution may lead to biased estimates depending on income level.

Next, we addressed the question of whether the willingness to pay (WTP) for a change in the probability of the occurrence of an adverse health effect differs between income groups. We showed that this was equivalent to asking whether the marginal rate of technical substitution between pollution and averting expenditures in the production of exposure differed between income groups. Out of the range available it would not be unreasonable to assume that the most efficient types of averting expenditures are undertaken first, and that furthermore less increases in averting expenditures are needed at higher levels of exposure to counteract an increase in pollution. These assumptions would imply that high-income groups have a higher WTP if they are less exposed at a given level of air pollution. WTP has empirically been found to be higher for individuals with higher income. This finding was contrasted with the fact that if the marginal utility of health decreases with higher health levels and if rich people have higher health levels, then the utility gain from an improvement in air quality is larger for lower income groups.

Having argued for the possibility that the negative health effects of air pollution are larger for lower income-groups and that the WTP to decrease the probability of illness
is larger for the higher income-groups, we then go on to address the four main income-related short-comings of "traditional" estimation practice and some improvements are suggested.

The first challenge is omitted variable bias. As long as behaviour affects the amount of exposure received, and as long as this behaviour may vary from one person to the next (depending on income), a regression of the adverse health variable onto only air pollution/exposure and a constant will give us a downwardly biased estimator of the direct effect of exposure on health. The reason is that an increase in exposure is followed by an increase in mitigating activity which counteracts the effect of exposure. Furthermore, a regression of the adverse health variable onto air pollution/exposure will yield an upwardly biased estimator of the total effect of exposure, i.e. of the dose-response coefficient, because high exposure may be associated with low income, which in turn may be associated with lower levels of mitigating activities and thereby higher adverse health effects. We can obtain unbiased estimates of the direct and total effect by including income and some measure of mitigating activity in the regression analysis.

Even if the slope of the dose-response function, as well as the direct effect of air pollution on adverse health are estimated correctly in one country, however, transferring either of them to a country with a different underlying income distribution will lead to wrongly valued health impacts. The problem in developing countries is often the lack of reliable, good quality data required to carry out the analysis. In the case where we want to transfer our estimated coefficient from a rich country to a developing country, we have to take into account that income, education, and baseline health levels will tend to be lower in the latter, and that the three are highly correlated. Although there is no clear-cut conclusion to be drawn as to whether the partial and total estimated effects of air pollution on adverse health would be higher in developing countries, several arguments seem to indicate that they may be. While waiting for higher-quality time-series data to be gathered in developing countries, a second-best approach to the problem was suggested. We suggested explaining variations in the regression coefficient between studies from different cities not only in terms of stochastic effects but also in terms of baseline factors or 'moderator variables' that influence the slope of the dose-response function. The coefficients derived from such a meta-analysis can then be used to transfer epidemiological results into contexts that differ in terms of these moderator variables.
We have further seen that even if the estimation of the slope of the dose-response function is correct, the quantification of the health impact will be wrong if the health effect of different income groups differs and if these income groups have different willingness to pay to avoid being ill. In particular, if low-income individuals tend to have a lower WTP and a larger direct adverse health effect from a dose of pollution, then low values of WTP will tend to be associated with high values of dose-response coefficients which in turn will imply a negative value for the covariance. In other words, the estimated health costs will tend to be higher than the actual costs. There are several approaches to transferring estimates of WTP from one country to another. The most common approach in the literature corrects only for income differentials between the countries, whereas a more sophisticated approach involves transferring a WTP-function using variables such as income, age, and chronic illness as moderator variables. In theory, similar approaches could be used to transfer estimates of WTP to different income sub-groups in one particular country (or age sub-groups), as we suggested in order to deal with the problem of the negative covariance and thus obtain better cost estimates.

The final problem addressed is that of bias due to errors in measuring the study variables. First, there may be serious errors in the measurements used of the adverse health effects. If low-income individuals 'choose' to work or to stay out of hospital during illness, then the traditional measurements of health effects of hospital days or lost work days will not capture the health effect on the poorer part of the population. The main impact is then a downward bias in the measured health effects from air pollution. The second and perhaps even more common measurement error occurs in measuring exposure to air pollution. We argue that for a given level of ambient pollution, exposure to it may be higher in low-income groups or countries, and estimated exposure-response relationship (assuming that we are measuring correctly the health effect) will be too low in higher-income countries as compared to low-income countries. Alternative measurement methods may be better able to capture both the morbidity effects of air pollution in the low-income groups and individual exposure, although each has its disadvantages.

There are at least three interesting insights/lessons to be retained from this discussion. Firstly, the failure to take the health effect of air pollution on different income groups properly into account may, as we have shown, lead to wrong estimates of the
health cost of air pollution, and therefore possibly to wrong decisions when carrying out benefit-cost analysis of pollution-decreasing policies.

We have further seen that the total effect of actual exposure on actual adverse health is likely to decrease in income because higher income groups have higher baseline health and are able to afford more mitigating expenditures. This, in turn, would imply that the common practice of transferring effect estimates from high-income to developing countries leads to under-valuation of the adverse health effects from pollution in these countries.

Although the actual effect of pollution on health may vary with income, and therefore make transfers of estimates from high-income countries to low-income countries particularly undesirable, carrying out studies in situ may however not necessarily lead to better estimates. If the effect of a certain change in pollution has a larger adverse health effect in low-income countries but less of it is captured by the adverse health effects measures, it is not clear whether the transferred estimate or the one estimated in the relevant country is more distorted. Following this thought one step further, increasing the income of the poorest groups may lead to less actual adverse health effects, but a larger amount of measured health effects.

With a larger adverse health effect from pollution on the poor part of the population, income-redistributing measures may decrease the negative health effect from pollution. Since income-redistribution possibly has positive effects on the growth rate of an economy (Benabou (1994), Galor and Zeira (1993)) this additional positive health effect may then lend support to a ‘double dividend’--type of argument for income redistribution.

Even in the absence of pollution reductions through pollution-targeting policies, the adverse health effect from pollution may possibly be reduced through such measures as the introduction of sick pay (especially relevant in developing countries), progressivity in the finance of health care, improvements in credit accessibility for the poor, or through income redistributing measures.

Crucial to developing efficient policies to decrease the health effects of air pollution in the future will be to understand the relative importance of exposure, averting and mitigating behaviour, and of baseline health for morbidity and mortality from air pollution, and to improve measurement methods.
PART 3: CAN POPULATION CHARACTERISTICS ACCOUNT FOR THE VARIATION IN HEALTH IMPACTS OF AIR POLLUTION?

A Meta-Analysis of PM$_{10}$-Mortality Studies.

6.1 Introduction

As early as in 1952, during the air pollution disaster in London,\textsuperscript{88} it was established that high levels of particulate-based smog could cause dramatic increases in daily mortality. The relationship between particulate matter and mortality has been analysed for some time now, and studies have reported evidence of increases in daily mortality also at much lower levels of particle concentrations. The variability among epidemiological findings, however, suggests that the connection between particulate matter and mortality is not well understood.

In this study we analyse the largest sample of short-term air pollution mortality studies to date, from the widest range of countries, in an attempt to further the understanding of the relationship between particles and mortality. In particular, our sample consists of time-series studies examining the effect of changes in daily (averaged) air pollution levels on daily mortality. The statistical relationship between particulate air pollution and mortality is addressed in epidemiologic studies, and the ensuing ‘dose-response functions’ tell us the impact on the mortality rate of a population of a certain dose of pollution.\textsuperscript{89} Because the epidemiologic studies differ in a number of ways, the regression coefficient of the dose-response function is likely to vary both with the characteristics of the exposed population, other site-specific differences, as well as analytical decisions.

This study will focus on whether population characteristics can explain some of the differences in effect estimates, while through sample selection trying to minimise the potential for other underlying sources for differences.

The analysis involves isolating relevant moderator variables using meta-regression methods. A moderator variable is a variable that causes differences in the correlation between two other variables, in this case between mortality and ambient concentration of air pollution. If there is true variation in results across studies, then one or more

\textsuperscript{88} The London smog disaster (December 1952) established that high levels of air-borne particles and sulphur dioxide produced large increases in daily death rates (HMSO, 1954).

\textsuperscript{89} For a review of the main study designs associated with epidemiologic studies refer to chapter 5.
moderator variables must exist that are able to account for the variance. The general underlying form is as follows:

$$b_j = \beta + \sum \alpha_k Z_{jk} + u_j \quad (j = 1, 2, ..., L) \quad (k = 1, 2, ..., M) \quad (6.1)$$

here $b_j$ is the reported dose-response estimate in the $j$th study from a total of $L$ studies, $\beta$ is the summary value of $b$, $Z_{jk}$ are the variables that could explain variations amongst the studies, $\alpha_k$ are the coefficients of the $M$ study characteristics that are controlled for, and $u_j$ is the error term.

The differences in results from individual studies imply that the quite commonly used procedure of transferring the regression coefficient unchanged to another population may lead to incorrect estimates of adverse health effects and the related costs. However, direct studies of the population in question may often not be feasible due to the quality of data, or to time and financial constraints. With the growing body of dose-response studies increasingly carried out also outside of the US, a second-best option is emerging. Rather than transferring the dose-response coefficients unaltered from one population to another, the existing studies can be used to estimate the coefficients on relevant moderator variables, and these in turn may enable us to transfer dose-response functions. We are then in the position to tailor-make the coefficient for local conditions.

The meta-analysis may hence serve three main purposes: it can increase our understanding of what affects the amount of deaths that are related to air pollution; it will help highlight areas where further studies may be needed; and finally, through the ensuing coefficients of the moderator variables, it may help transferring the dose-response coefficients to countries where empirical studies have not yet been feasible or to forecast the effects of policies targeting air pollution.

The meta-analysis in this paper will consist of two stages. First, the Maddison and Gaarder (2001) study will be re-evaluated. Although an interesting first approach to the subject of moderator variables, we argue that one of the observations included in the MG-sample should have been excluded, that the regression method used is implicitly based on an unrealistic assumption, and that potentially relevant variables were excluded in the analysis and may have biased other variable coefficients. The changes in the sample, moderator variables, and methodology will change the results dramatically. Second, a meta-analysis based on a larger sample of observations (studies) and moderator variables is carried out and

---

90 Button and Nijkamp (1997).
the ensuing results are presented. For a review of how the present study relates to the Maddison and Gaarder study, please refer to appendix 6.A.

Section 6.2 introduces the concept of meta-analysis, as well as the various uses, strengths, and weaknesses of this type of analysis. In section 6.3 we then move on to presenting the moderator variables selected. Next, section 6.4 describes the criteria used in composing our sample of past studies, the data used to capture the moderator variables, as well as the model and estimation procedure. The main results are presented in section 6.5, together with a sensitivity analysis and a discussion of the findings. Main results and implications are summarised in chapter 6.6.

6.2 A Survey of Existing Meta-Analyses of the Mortality from Air Pollution

Meta-analysis involves the synthesising of previous empirical analyses. Before presenting a survey of what has been done in this field, a brief introduction to the original studies and study-designs upon which these meta-analyses are based is therefore required.

The majority of the original studies use time-series data to examine the effect of short term responses in mortality to changes in air pollution levels. The main advantage of time series studies over cross-sectional studies is that socio-economic and demographic characteristics of the population are unlikely to change and do therefore not require explicit modelling. The studies usually assume that the daily death counts \( Y_i \) are Poisson-distributed\(^1\) with:

\[
\log (E(Y_i)) = X_i \beta
\]

where \( X_i \) is the vector of covariates on day \( i \), \( \beta \) is the vector of regression coefficients, and \( E \) denotes expected value. The unit of analysis in these studies is the day, and hence the potential confounders that must be controlled for are those that vary over time, possibly in coincidence with air pollution. Based on this logic, the vector of explanatory variables typically contains terms corresponding to a measure of air particulate, as well as meteorological covariates (e.g. ambient temperature and relative humidity), long-term and

---

\(^1\) Only a small portion of a population dies on any given day. The number that die is a count; i.e. it can only take on values limited to the non-negative integers. This suggests that a Poisson process is the underlying mechanism modelled, since in a Poisson process a homogeneous risk to the underlying population is assumed. Given that underlying risk, the probability of \( Y \) deaths occurring on a given day is given by:

\[
\text{prob}(y; \lambda) = \frac{e^{-\lambda} \lambda^y}{y!}
\]

where \( \lambda \) is the expected number of deaths on any day (i.e. \( E(Y) \)) (Schwartz et al. (1996b)).
seasonal trend components, disease epidemics (e.g. influenza episodes), and day of the week and holidays.

It is important to point out that the dose-response function technique, as presented above, is mechanistic, incorporating no model of how individuals behave. The dose-response coefficient relies on the socio-economic and demographic characteristics remaining unchanged. Although it can quite accurately describe the effect of a change in air pollution on mortality in a certain population, demographically different groups and groups subject to different economic constraints may respond differently to exposure to air pollution. This is why, when we compare results from studies carried out at different sites, we need to take such differences into account. That is the role of the moderator variables in the meta-analyses.

Early meta-analyses were mainly concerned with finding the average effect across studies, implicitly assuming that the estimated effect in each study is an estimate of an effect size common for the whole population of studies. More recently, meta-analytic work has started to focus upon discovering and explaining the variations in effect sizes (Raudenbush and Bryk (1985)).

Button and Nijkamp (1997) discuss a number of issue areas within environmental policy evaluation which could benefit from the use of the meta-analysis techniques. In evaluating environmental costs, the meta-analysis can be used to look for indicators of central tendency in previous case studies or, alternatively, to explain why the studies generate differing results. Furthermore, meta-analysis can be used in connection with the assessment of the effectiveness of alternative policy instruments in containing environmental damage, the assessment of political acceptability of alternative environmental instruments by decision makers, exploration of the appropriate political level of intervention to contain environmental damage, and finally in forecasting the effects of environmental policies.

Rosenthal (1991) distinguished three purposes of meta-analyses. First, to summarise for a set of studies what the overall relationship is between two variables investigated in each study. Second, to look at the factors associated with variations in the nature of relationships between two variables over a range of studies. Finally, to look at the aggregate data for each study and correlate this with other characteristics of the study (Bergh et al. (1997)).
There is a wide range of problems involved in employing meta-analysis in economic research. Broadly, we can divide the problems into two categories. The first category has to do with the objectivity with which the information is collected and reported, whereas the second deals with comparability between studies and how well the studies are designed for the particular question they want to address. There is a possible bias resulting from the nature of the studies that are included or excluded. First, the researchers use various inclusion-selection rules for the analysis (e.g. including only published studies) which are inherently subjective. Second, there is the tendency to publish only positive results. As for comparability, a number of challenges exist. Studies often use diverse units of output measures and, furthermore, diverse methods of obtaining these outputs (e.g. diverse regression methods, different sets of control variables). A degree of subjectivity is introduced into many of these studies and thereby into the meta-analysis because the reported results were based on what, in the authors' opinions were the best coefficient estimates obtained. In particular, some studies reported coefficients obtained using same day level of pollution, others used one-day lags, and others again used moving averages of different lengths.

Estimates can differ partly due to the fact that the studies use different samples of the total population and partly due to the differing conditions under which the research takes place. Fixed effects models assume the existence of a common effect size in all the studies, whereas random effects models assume a different real effect in each study. In the latter case, combining effect sizes from empirical studies means assessing the average size of the real effect.

If we reject the hypothesis of equal real effect sizes, the next question is then whether we can find moderator variables that explain the variations between the empirically estimated effect sizes. If a linear combination of variables exists that completely explains the variations in the real effect sizes, then the effect size is fixed and not random (although the real effect sizes are different in each study). This is, however, a rare case. In most cases it is more realistic to use a model that takes account of the imperfections of the explanatory model.

Before we review the literature of meta-analyses on studies of mortality from particulate matter it is useful to understand the various models and assumptions underlying the different approaches.
Let us assume that the estimated effect size $d_i$ of study $i$ is equivalent to a true effect size $\delta_i$ plus an error of estimate $e_i$, where the errors are assumed to be independent and normally distributed with a variance $\nu_i$:

$$d_i = \delta_i + e_i, \quad i = 1, \ldots, k, \quad e_i \sim N(0, \nu_i)$$

(6.2)

The random effects model assumes that the effect size parameters $\delta_i$ can be decomposed into a mean population effect $\theta$ and a between-study variability term $u_i$, where the errors are assumed to be independent and normally distributed:

$$\delta_i = \theta + u_i, \quad i = 1, \ldots, k, \quad u_i \sim N(0, \tau^2)$$

(6.3)

The mixed effects model (which is equivalent to a random effects model incorporating study characteristics) assumes that the effect size parameter is a function of known study characteristics and random error:

$$\delta_i = W_i' \gamma + u_i, \quad i = 1, \ldots, k, \quad u_i \sim N(0, \tau^2)$$

(6.4)

The fixed effect model implicitly assumes no between-study variability in either of the equations above, i.e. $u_i = 0$.

Combining equations (6.3) and (6.1) we obtain:

$$d_i = W_i' \gamma + u_i + e_i$$

(6.5)

Therefore, assuming that the error terms are independent, we can express the marginal distribution of $d_i$ from the mixed effect model as:

$$d_i \sim N(W_i' \gamma, \nu_i + \tau^2)$$

(6.6a)

The distribution of $d_i$ from the random effect model, the simple fixed effect model, and the fixed effect model with study characteristics can be expressed respectively as:

$$d_i \sim N(\theta, \nu_i + \tau^2), \quad (6.6b) \quad d_i \sim N(\theta, \nu_i), \quad \text{and} \quad (6.6d) \quad d_i \sim N(W_i' \gamma, \nu_i)$$

Based on the epidemiological literature dealing with the relationship between air pollution and mortality, to our knowledge seven meta-analyses have been carried out: Ostro (1993), Schwartz (1994), Lipfert and Wyzga (1995), Environmental Protection Agency (1996), Levy et al. (2000), Institute for Environmental Studies (2000), and Maddison and Gaarder (2001).

After converting the results of the different studies into a common metric, Ostro’s meta-analysis derived the unweighted average of central estimates and found that the mean effect of a $10 \mu g/m^3$ (micrograms per cubic meter) change in PM$_{10}$ on the percentage change in mortality varied between 0.64 and 1.49 per cent. Lipfert and Wyzga, on the other hand, calculated the variance weighted average of air-pollution-
mortality elasticities and found that the mean overall elasticity as obtained from time-series studies for mortality with respect to various air pollutants entered jointly was approximately 0.048 (0.01 – 0.12). The elasticity obtained for population-based cross-sectional studies was of similar magnitude. The models used in both of these meta-analyses implicitly assume that each coefficient estimate, $\beta$, is a random sample from a single underlying distribution with a distribution as in expression (6.6c). Ostro’s study in addition implicitly assumes equal estimation errors in all of the studies.

Joel Schwartz (1994c) carried out a meta-analysis on a set consisting of studies from the US, London, and Athens. The main aim of the analysis was to compare the results found in different studies to the levels of potential confounders and the correlation between particulate matter and potential confounders in the individual studies to assess the likelihood that the results are driven by inadequate control for those factors. It then combines the studies in a meta-analysis and computes the average percentage increase in mortality per unit of pollution. Three approaches to calculating this average were used; unweighted, variance weighted, and quality weighted. The latter weights were based on the possibility in each study that the true effect sizes vary at least in part as a function of multiple identifiable sources, or confounding variables, and that if these have not been taken properly into account in the regression model used in a particular study the random error term will be larger for these studies. The central concern in the study was of confounding by some other pollutant, by weather and season, and an additional concern was the quality of the exposure assessment. Studies were given a higher weight the more they controlled for confounding factors (the highest weight was 4 and the lowest 2). The unweighted meta-analysis, as well as the analyses using the various weighting options all gave a relative risk of 1.06 for a 100 $\mu g/m^3$ increase in total suspended particulate mass, which implies that the relationship is highly unlikely to be due to confounding factors.\(^\text{92}\)

By introducing the quality weights, Schwartz is allowing for the idea that there is no single common underlying effect size. However, the size of the weights was provided by the researcher based on his subjective opinion of the quality of control for confounding. This subjective weighting may influence the results and is a weakness of the above meta-analysis.

---

\(^\text{92}\) The information concerning the effect from exposure to air pollution on the risk of mortality uncovered by regression analysis can be expressed in a number of alternative ways. The findings are often expressed in terms of relative risk. The relative risk indicates the ratio of the probability of occurrence of a given effect between two different exposure levels or exposure groups.
In its meta-analysis the U.S. Environmental Protection Agency (EPA (1996)) criteria document used a random effects model to estimate PM mortality, where the distribution of the effect parameter is assumed to be given by expression (6.6b). The relative risk for mortality from PM$_{10}$ exposure averaged over 2 days or less was in this study estimated as 1.031 per 50 µg/m$^3$ PM$_{10}$ (CI: 1.025 – 1.038), whereas for a longer averaging time of between 3 and 5 days the relative risk was estimated as 1.064 (CI: 1.047 – 1.082). When potential confounding pollutants were included in the model the relative risk estimate decreased (1.018, CI: 1.007 – 1.029). Although the random effects model can quantify the amount of residual variance that can be explained by study characteristics, it does not attempt to identify what these characteristics are or how they influence the effect estimates.

In the most recent meta-analysis carried out by the Institute for Environmental Studies (IVM) (2000), the purpose was to obtain a single pooled estimate of the health effects reported from the selected studies in order to use this for evaluating the benefits gained from improving air quality in Mexico City. A weighted average was computed, giving more emphasis to studies with lower error in estimating their regression coefficient, as well as studies carried out in Mexico City (‘articles with estimates based on Mexico City were given double the weight of international cases, because they are more likely to reflect the socio-demographic and susceptibility characteristics of the Mexico City population’ (p.27, IVM (2000))). The pooled estimate of the effect of PM$_{10}$ on total mortality was 0.79 per cent change per 10 µg/m$^3$ daily average PM$_{10}$ (CI: 0.06 – 1.68). There is a certain inconsistency/contradiction in the method they have used. By weighting the estimates according to the inverse of their variance, the study is assuming that the variability in reported effects is attributable solely to sampling error. On the other hand, giving higher weights to the studies carried out in Mexico City implies that the authors assume that these studies are capturing some local characteristics and are hence more relevant for the purpose of policy-evaluation in Mexico City. This is a rather indirect way of controlling for confounding factors and may weaken the reliability of the pooled estimates. Furthermore, as was the objection to Schwartz’ study, the size of the weights was provided by the researcher on a rather ad hoc basis and may influence the results.

Rather than providing pooled effect estimates, the meta-analysis by Levy et al. (2000) addresses between-study variability potentially associated with analytical models,
pollution patterns, and exposed populations. They use the mixed effects Empirical Bayes (EB) model derived by Raudenbush and Bryk (1985), assuming that variability is due partly to sampling errors (or intra-study variability) and partly to between-study variability. This method is used in the present study as well, and the details of the method are set out in section 6.4.3. With a sample of 29 observations, 19 from the United States and 10 from outside of the United States, they investigate whether the ratio of PM$_{2.5}$ to PM$_{10}$, other pollutants, climate, season, prevalence of gas stoves and/or central air conditioning, percentage of elderly, percentage in poverty, and the rate of mortality can explain some of the differences in effect estimates. When analysing the 19 PM studies from the U.S. for which more confounding variables were available, the mortality rate was estimated to increase by 0.7 per cent per 10 μg/m$^3$ increase in PM$_{10}$ concentrations, with greater effects at sites with higher PM$_{2.5}$/PM$_{10}$ ratios, supporting the hypothesised role of fine particles. When all of the 29 studies were included, but only a subset of the predictors were available (PM$_{10}$ concentration, averaging time and lag time, percentage of the population older than 65 years of age, baseline mortality rate, heating and cooling degree days, and dummy variables for PM$_{10}$/TSP and U.S./non-U.S. studies) only baseline mortality rate was significant. The grand mean estimate was about the same as for the 19-studies sample.

Finally, the meta-analysis by Maddison and Gaarder (2001) investigates whether, in a sample of 13 European and developing country studies, some of the between-study variability can be associated with pollution levels, the percentage of the population over 65 years of age, average income level, and the level of income inequality at a certain average income level. By weighting the effect estimates according to their estimated variances, we implicitly assumed a fixed effect model with study characteristics for which the distribution was given in expression (6.6d). The study found that the effect estimates were significantly affected by the percentage of the population over 65 years of age, as well as income distribution. Based on the data used in our study, a model without predictors (i.e. fixed effect model) gives an estimate of the effect of PM$_{10}$ on total mortality of 0.3 per cent change per 10 μg/m$^3$. An implicit assumption of our analysis, which seems unlikely and therefore weakens the results of this study, is that all the variance among the study effects other than sampling variance can be explained as a function of the study characteristics we chose to include.
In addition to the meta-analyses discussed above, a number of review articles have relied on qualitative discussions of the credibility of the evidence related to potential confounding factors (e.g. climate, correlated pollutants). Some of the authors of these studies conclude that a causal relationship clearly exists (Brunekreef et al. (1995), Pope et al. (1995a), Pope et al. (1995b), Thurston (1996)), whereas others (Gamble and Lewis (1996), Moolgavkar and Luebeck (1996)) argue that the relationship is spurious. The lack of quantitative base, however, makes these review studies more vulnerable to the set of studies chosen and the points the authors wish to argue. (Levy et al. (2000)).

6.3 The Moderator Variables

As mentioned in section 6.2, the original studies do not explicitly model the demographic and socio-economic characteristics of the population studied. The reason for this is that for a factor to confound the relationship between pollution and daily mortality it must be correlated with both pollution and mortality. Therefore, characteristics such as baseline health, age, and income cannot induce an association between today’s mortality count and yesterday’s air pollution, since they are not correlated with air pollution and do not vary on a daily basis. For cross-sectional mortality studies, on the other hand, personal characteristics and habits are important potential confounders, whereas short-term weather changes are not (Schwartz (1994c)). In our meta-analysis the aim is to combine time-series studies cross-sectionally, and to explain the variation in the dose-response coefficients using moderator variables (also known as effect modifiers). These moderators will hence need to address cross-sectional differences, rather than factors changing over time.

When deciding on which study-characteristics to include as potential predictors or moderator variables three factors guided the selection; theoretical plausibility, availability of characteristic-data, and novelty. This led to the following moderator variables; mean particle levels, amount of elderly people in the population, income level, income distribution, education, baseline health, and health services. The reasons why we believe these factors (or characteristics) to be potential moderators are discussed below. Other study characteristics, such as the lag and averaging times, the levels of other pollutants, the ratio of fine particles to overall particle concentration, and the type of

---

93 The most important confounders for the relationship between air pollution and daily mortality are weather and infectious disease epidemics, according to Schwartz (1994c).
94 The first four were also used in the Maddison and Gaarder meta-analysis.
mortality considered, although potentially interesting predictors, were either not considered due to lack of information in many of the studies or were investigated through sensitivity analysis.

Most dose-response analyses have implicitly assumed a log-linear relationship between the mortality count and pollution, however, it has been argued that this may not be accurate. As the exact shape of the relationship is not yet known, we argue that it may be interesting to include pollution as a moderator variable. By regressing the estimated pollution-mortality association on pollution (i.e. second order partial derivative), we pick up any non-linearities in the relationship.

A variety of advanced disease states, as well as generally lower baseline health levels, may predispose individuals to heightened susceptibility to premature death due to exposure to air pollution. This implies that the death rate due to a certain amount of particle exposure may increase more among elderly and individuals with lower baseline health as compared to the younger and those with better health, and that death rates due to respiratory and cardiovascular failure increase more than the total rate. However, as exposure tends to be approximated by air pollution concentration measurements from central monitoring stations, it is possible that the individual exposure for a certain amount of pollution concentration also varies with baseline health levels and age (i.e. the optimal amount of averting activities may be affected by age and health level). Furthermore, the heightened susceptibility to exposure may influence the amount of mitigating activities chosen by elderly individuals and individuals with low levels of baseline health. In other words, both age and baseline health levels may well influence the amount of mitigating and averting activities undertaken, and hence affect health indirectly. On the one hand, it is possible that an individual with low health levels will be more inclined towards trying to prevent further adverse health effects (both due to personal experience with bad health and due to decreasing utility at an increasing rate). On the other hand, the individual may be used to being in bad health and expect to live for a shorter time, and therefore less inclined to invest in health. It is theoretically not clear what net effect baseline health and age will have on the concentration-response coefficient, but both characteristics may certainly play a role and should therefore be included as moderator variables.

Empirical dose-response studies have found that mortality among the elderly is more responsive to changes in particulate pollution than is mortality for the entire population or mortality among the younger generation (Ostro et al. (1996), Schwartz and
Dockery (1992a). Evidence further suggests that air pollution has its greatest adverse
effects on people with pre-existing chronic conditions such as asthma, bronchitis, and
emphysema (Ostro (1987)).

Age and baseline health will tend to be closely associated when looking at entire
populations. In particular, if a population has a large percentage of elderly people it
indicates that the baseline health of that population is rather high, enabling so many to
live to an old age. Hence, if the baseline health variable is omitted from the regression
analysis the age-variable, which is supposed to pick up the part of the population that is
most at risk from high air pollution, will also proxy for the average health level of the
population. These are two offsetting effects, and age will hence tend to be biased
downwards. Baseline health levels will tend to be associated with level of income,
although the association will probably be highly sensitive to the measure used for
baseline health. Low-income individuals may have worse baseline health levels if low
income and little education have given rise to wrong and/or insufficient nutrition and
other health investments in the past. On the other hand, people who have a history of
chronic obstructive pulmonary disease or cardio-pulmonary problems are also thought to
be particularly vulnerable, and these types of health problems are more pronounced in
high-income groups and countries.

There are several reasons why one would expect the increase in mortality due to
ambient particles to vary with income, and these were thoroughly discussed in chapter 5.
Firstly, for a certain increase in ambient concentration of air pollution we argued that
lower income groups were likely to experience a larger increase in exposure than were
the higher income groups because the former are not being able to afford much averting
activities (e.g. sealing houses to reduce the penetration of outdoor pollutants, using less-
polluting heating and cooking fuels, spending less time in traffic). Secondly, for a certain
amount of exposure and its anticipated health effect we suggested that the behavioural
response (e.g. visiting a doctor, taking medication) will typically be influenced by income
level. These mitigating measures imply costs which poor people may not be able to
afford, or willing to pay given their budget constraints. Finally, we argued that there may
be differences between low and high income groups, and even more so between low and
high-income countries, in the extent to which official mortality statistics reflect actual
mortality. It is not unlikely that deaths among the poor will be underrepresented or

---

^5 Refer to section 5.4.3 in chapter 5.
unavailable in official statistics. Although this latter point may imply an under-representation of the increase in mortality due to air pollution in lower income groups or countries, we suggest that the overall measured adverse health effects of an increase in air pollution will tend to be larger in low-income countries than in higher-income countries. Income should therefore be included as a mediator variable.

There are additional reasons why exposure may differ between developed and developing countries, and why an increase in exposure may lead to a larger increase in mortality in low than high-income countries that are not necessarily due to income levels, although income may be part of the underlying explanation for these factors. Firstly, the effect of an increase in pollution on exposure may be larger in low than high-income countries due to the fact that many low-income countries are situated in warm climates and the residents in these climates are therefore likely to spend a greater portion of their time outdoors (Ostro (1994)). Other differences between low and high income countries may also influence the amount of time spent outdoors, such as crime rates, indoor air pollution, and social interaction traditions. Furthermore, the pollution level locally at the work place may be higher in less developed countries due both to the cost of abatement and less strict work place regulation. Finally, an increase in exposure may lead to a larger increase in mortality in low than high-income countries due to the quality and availability of health care. In addition to the often very restricted availability of health care, the quality of health care in developing countries is often poor, something which may affect the efficiency of mitigating measures. Hence, the risk of dying from the health effect of air pollution may be influenced by own behaviour or by the facilities available, and could be higher for lower income groups or cities. Due to the lack of reliable data on availability and quality of health care, time spent outdoors, work place pollution etc., such variables have in general not been included as moderator variables. By excluding these from our analysis, we implicitly allow income to proxy for their effects.

There are at least three reasons to believe that the income distribution in a country, i.e. relative income, is important for the difference in health effects. First, unless the effect of income on the dose-response coefficient is linear, using an average income variable will not capture correctly the sum of the effects of each individual's income on his or her adverse health. In particular, there are probably decreasing returns to averting and mitigating activities which would imply a tendency for higher inequality to be
associated with higher mortality rates. Second, the location at which people live within a city will affect the amount of air pollution they are exposed to. Although individuals can move between cities, it seems likely that for most cities housing prices are determined by within-city demand. Hence, it is not so much the income level as the position within the income distribution that determines where an individual lives. If individuals with relatively low income tend to live in the most polluted areas, as evidence suggests, and if the adverse effect of air pollution is larger on lower income individuals (due to lower baseline health, less education etc.), then this would once again imply a larger PM$_{10}$-mortality in cities with large income-inequality. Third, there is a line of research that implicates the biochemical effects of psychological stress as a risk factor, and relates this stress to social status (Deaton and Paxson (1999)). Social status can then be modelled as income relative to the average income. If mortality is associated with stress, and stress is related to social status (income relative to the average income), then this is a third reason why higher income inequality may lead to a larger mortality rate from air pollution. GINI may be proxying for inequalities in baseline health or for the quality and availability of health care, if satisfactory measures for these two variables are not available.

It may also be of interest to consider whether the effect of income inequality on the mortality rate from air pollution varies according to the average level of income at which the inequality takes place. On the one hand, one could speculate that high income inequality in a low-income country would imply a large amount of people not being able to undertake any averting and mitigating activities whatsoever (demand-side), and that only a small increase in income and health investment for these population groups therefore would have a large effect in reducing mortality. On the other hand, the range of averting and mitigating measures available to the public (the supply-side) and the information about the effects of pollution and how to minimise these may well be larger in high-income countries, implying that the way in which income is distributed may play a more significant role in determining the amount of deaths caused by particulates.

---

96 There are several reasons why we find decreasing returns to health investment likely. First, it is reasonable to assume that the most cost-efficient measures are undertaken first. Second, it is not unlikely that a similar health investment measure has a larger positive effect at high levels of exposure and low levels of baseline health than at lower levels of exposure and better health levels, and the two latter characteristics tend arguably to be associated with higher income groups.

97 A cautionary remark is in order: if the effect of air particles on mortality were to be increasing at a decreasing rate, then the above finding would not necessarily hold. Empirical evidence so far, however, does not support this, rather, a linear or even a convex function are usually assumed.
Furthermore, high income-inequality in high-income countries may arguably lead to more psychological stress than in low-income countries. For the above reasons, we suggest considering both the effect of a relative income-inequality measure, as well as an income distribution measure that takes the average level of income into account as moderator variables in the regression analysis.

The income variable we ideally would like to have is the average income in the location in question (be it a city or otherwise), or even more precisely, the average income of the vulnerable population within the relevant location. We were not able to obtain this information, however, and had to settle for a second-best option, namely the average income in the country in question. If the average level of income is similar in the study location as it is in the country as a whole, this will be a satisfactory approach. However, it is not unlikely that for the study locations, most of which are relatively large cities, this will not be the case. If the average income level in large cities differs in a consistent manner from the country average, then it is possible that the best way of capturing the average city-income is a composite of the average country income and the income distribution in the country.

The level of education may affect the knowledge people have about health, health production, and the connection between air pollution and health, and hence affect the level of baseline health, as well as the amount of averting and mitigating expenditures undertaken and the efficiency of these expenditures. Although there is conflicting evidence as to whether little knowledge/education leads to over or under-investment in health, we suggest that education should be included as a moderator variable. Since schooling is closely associated with income, the income coefficient will probably proxy for this variable if it is not included as a moderator variable.

Finally, the health services provided in a country are likely to influence the amount and severity of adverse health incidences. Health services are likely to be highly positively correlated with baseline health and the amount of people over the age of 65, and negatively with income inequality.

The expected signs of the moderator variables presented and discussed in this section are summarised in table 6.1 below.
Table 6.1: Summary of the expected signs of the moderator variables in a table

<table>
<thead>
<tr>
<th>Moderator variable</th>
<th>Sign</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air Pollution</td>
<td>+</td>
</tr>
<tr>
<td>Baseline Health</td>
<td>-</td>
</tr>
<tr>
<td>Age</td>
<td>+</td>
</tr>
<tr>
<td>Income</td>
<td>-</td>
</tr>
<tr>
<td>Income Inequality</td>
<td>+</td>
</tr>
<tr>
<td>Composite Variable (Interaction Variable) of Income and Income Distribution</td>
<td>?</td>
</tr>
<tr>
<td>Education</td>
<td>-</td>
</tr>
<tr>
<td>Health Services</td>
<td>-</td>
</tr>
</tbody>
</table>

All of the above mentioned variables are potentially important moderator variables, especially when transferring estimates to cities in developing countries which may take substantially different values on all of these. The level of air pollution is in general significantly higher in many developing countries than in the developed countries that generated most of the literature. Furthermore, an important difference between developed and developing countries is that the former tend to have an ageing population, whereas the latter have a majority of young people (higher birth-rate and lower life-expectancy), and we therefore find it potentially interesting to include this moderator variable. A crucial difference between developed and developing countries is the lower average income level in the latter. In addition, income in developing countries tends to be more unequally distributed (the average GINI-coefficient for the low-income countries in the World Development Report 1998/99 is 0.41, and for the high-income countries it is 0.30). As for health levels and education, these are both closely associated with income and thus typically on average much lower in developing countries than in their richer counterparts.

If the original studies have not satisfactorily controlled for confounding variables such as other pollutants, the ratio of fine particles to overall particle concentration, temperature, season, and humidity, and if these are correlated with the measured ambient particles, the resulting dose-response coefficients may be biased. However, assuming that the original studies have (linearly) controlled for various confounding factors, these may still have an impact on the measured effect size of air pollution on mortality. As information on these variables was missing in many of the studies in our sample such moderator variables were left out. However, a meta-analysis focusing on such confounding variables was carried out by Levy et al. (2000).
6.4 Sample Selection, Data, and Methodology

6.4.1 Sample

The first sample on which the restricted meta-analysis is performed below is based on the sample used in a paper by Maddison and Gaarder (2001) and will be referred to as the MG sample. This was mainly based on 7 time-series studies (10 observations) resulting from the APHEA project (see Katsouyanni (1997) for an overview). The sample was supplemented with studies from Chile (Ostro et al. (1996)), Sao Paolo (Saldiva et al. (1995) and Delhi (Cropper et al. (1997)). As compared to the original study, one study has been excluded; namely the one on Sao Paolo because it consisted of the over-65 year-olds only.

The criteria for inclusion in MG paper were the following: i) papers including the quantification of either Total Suspended Particles (TSP), Black Smoke (BS), or any Particulate Matter (PM); ii) published papers evaluating the association between exposure to particles and total non-accidental mortality; iii) mortality figures modelled using Poisson regression analysis, and; iv) studies resulting from the APHEA-project or studies carried out in developing countries. Papers that did not present information on the variance, standard error, or confidence intervals of the estimated coefficient were excluded, as were studies which did not control for temperature and seasonal variation over the study time period.

One of the studies included in the sample (Sao Paolo, Brazil, by Saldiva et al. (1995)) looked only at the effect of air pollution on individuals over 65 years of age. There are several reasons why it is wrong or undesirable to include the Sao Paolo study in the analysis.

First, the data used to capture the coefficients of the moderator variables are based on the overall population of the country, and clearly the sample based on the part of the population which is over 65 is not representative of the population as a whole. In particular, there is no reason to expect the level of average income to be the same in this study sample as in the overall population, nor the distribution of income. Furthermore, it is highly unlikely that the average amount of education and baseline health should be the same in the group of over 65 individuals as in the population as a whole.

98 For more details about the sample refer to Maddison and Gaarder (2001).
Second, even if the correct data to capture the value of the moderator variables for the sample of over 65 were available, including this study in the analysis is still problematic. Unless we are certain that we have exhausted all possible reasons for differences in estimated slope between the studies by including the relevant moderator variables, the results may well be spurious. The fixed effect (captured by $\gamma_0$ in equation (6.8) below and by $\beta$ in equation (6.1)) may be higher in Sao Paolo or Brazil as a whole compared with other countries, but this we cannot know as we only have information on the sample of people over 65 years of age. Since the high dose-response coefficient coincides with a high value for the age-variable, this variable receives a large significance, but it could well be that there is no age-effect and just an unusually large fixed (unexplained) effect. There is no way of knowing without a more representative sample.

Finally, regressing mortality rate ($h$) on a constant ($f$) and the level of ambient pollution ($p$) for different age groups we expect to find the constant to be higher for elderly people. Since we are concerned with non-accidental deaths, we can assume that the constant is increasing with age, $a$ (at least if we exclude infant mortality). Let us furthermore assume that the effect of a change in pollution on the mortality rate is also increasing in age. In simplified terms we can then write the regression equation as follows:

$$h_i = f_i(a_i) + \beta_i(a_i)p + \epsilon \quad f_o>0, \quad \beta_o>0$$

(6.7)

In the present meta-analysis we regress the dose-response coefficient that resulted from the study onto, among others, a variable picking up the age-factor (per cent of the sample over 65):

$$b_i = \gamma_0 + \gamma_i a_i + u_i$$

(6.8)

When regressing the dose-response coefficients on the age variable, the fact that age affects the constant in the original dose-response function is not taken into account. The coefficient on the age-variable will hence be biased upwards:

$$\gamma = \frac{\partial \beta_i}{\partial a_i} = \frac{\partial(h_i(a_i) - f_i(a_i) - \epsilon)}{\partial a_i} \frac{\partial f_i}{\partial a_i} < \frac{\partial h_i}{\partial a_i} / p$$

Turning to the larger sample on which we perform a meta-analysis, time-series studies for this analysis were gathered from previously published meta-analyses or
review articles (Maddison and Gaarder (2001), Levy et al. (2000), Institute for Environmental Studies (2000)), as well as from PubMed.99

The selection of the wider sample is based on the following criteria for inclusion:

1. papers including the quantification of either Total Suspended Particles (TSP), Black Smoke (BS), or Particulate Matter (PM) larger than 2.5 μm in diameter;
2. published papers evaluating the association between exposure to particles and total mortality;
3. mortality figures modelled using Poisson regression analysis;
4. studies carried out on a representative sample of the population (e.g. excluding studies carried out on particular age groups); and
5. analysis controlling the confounding effect due to meteorology and temporal effects.

Papers not presenting information on the variance, standard error, or confidence intervals of the estimated coefficient were excluded. Furthermore, papers reanalysing the same site and time period (either by the same or different authors) were excluded on the grounds of double counting. Instead of restricting the sample to APHEA and any available developing country studies, all available studies were included. In total, 70 estimates from 56 studies and 21 countries were selected.

A number of factors potentially influencing the estimated dose-response coefficients were not used as criteria for inclusion or exclusion, but were rather the subjects of sensitivity analyses. In the case of the total mortality measure, we found it interesting to investigate whether inclusion of studies looking at all-cause mortality rather than just non-accidental mortality had a significant effect on the regression results. Similarly, testing the sensitivity of our findings to the air particle measurements used, as well as the lag structure, could potentially yield new insights into the underlying relationship between air pollution and mortality.

A further factor likely to affect the estimated association between exposure and health in low and high-income countries differently is the way in which exposure has been measured. As adequate information on indoor air pollution in different countries was not available this factor could not be subjected to a sensitivity analysis, however, it will be important to keep in mind when interpreting our results. Ambient pollution at central monitoring stations may be particularly ill-suited to capture particulates exposure

99 PubMed, a service of the National Library of Medicine, provides access to over 11 million citations from MEDLINE and additional life science journals.
in low-income countries. Studies have found that indoor air pollution levels are as high if not higher than outdoor levels in several developing countries due to lack of air conditioning and some indoor sources present (e.g. Chestnut et al. (1998), Baek et al. (1997)). If it is vulnerable people (low baseline health levels, or of higher age) who tend to die from air pollution, and if indoor air pollution does not strongly covary with outdoor air pollution from day to day, then the exposure-response association may be much larger but not be captured by studies that use readings from central monitoring stations to measure exposure. In other words, those who are vulnerable to outdoor air pollution may already have died from indoor air pollution.100

As for the amount of pollutants included in the regression model, it could be used neither as inclusion/exclusion criteria, nor as a subject for sensitivity analysis. The main reason for this is that many studies were unclear as to whether the final results they reported for the particulate mortality coefficient were actually based on single, dual, or multiple pollution models. From the studies that did express clearly the amount of pollutants involved in their regressions we know, however, that a large majority of the time-series studies included in the MG-sample and full sample of this paper feature single-pollutant rather than multi-pollutant regressions. The potential drawbacks of both single and multiple pollutant regressions are discussed briefly below.

Some epidemiologists are uneasy with the reliance on single pollutant regressions because different pollutants tend to be highly correlated (Moolgavkar et al. (1995)). They argue that it is premature to single out one of them as being responsible for the observed correlation between air pollution and mortality. Furthermore, the use of single pollutant models renders the interpretation of the available evidence difficult, since it is not known if the deaths attributed to the different air pollutants are additive or not. Finally, choosing one pollutant as a marker for air pollution can lead to under-estimation of the problem if in fact several air pollutants are responsible.

The use of single pollutant regressions has been defended in the literature by Schwarz et al. (1996b). They argue that given the correlation between the pollutant variables

100 Studies (e.g. Baek et al. (1997), Chestnut et al. (1998), Janssen et al. (1998)) comparing indoor and outdoor concentrations of air pollution found the difference to be attributable in part to human indoor activities (e.g. type of stove used for cooking and heating, ventilation, tobacco smoke). Clearly, the more indoor air pollution is attributable to indoor activities, the less indoor air pollution will covary with outdoor air pollution.
and the relatively low explanatory power of air pollution for mortality, including multiple pollutants in the regression risks letting the noise in the data choose the pollutant.

We will assume that the studies selected on the basis of our selection criteria were independent samples from a random distribution of the conceivable population of studies. In section 6.4 we will return to this issue and discuss why this assumption may be difficult to support.

6.4.2 Data

A number of airborne particulate measurement methods have been used in exposure-response studies. Gravimetric (weight) measurements of collected particles yield direct measurements of airborne particle mass. The high-volume sampler collects and measures the mass of total suspended particulates (TSP), whereas more recent samplers include devices to selectively collect and measure the mass of various size fractions of PM (e.g. PM\textsubscript{10}, PM\textsubscript{13}, PM\textsubscript{2.5}). Two optical, and thus indirect, methods of measuring the mass of collected particles have also been frequently used. The black smoke (BS) method is based on light reflectance from particle stains on sample collection filters, whereas the coefficient of haze (COH) method is based on light transmission through the filter stain. According to the EPA, credible estimates of particle concentrations (in \(\mu g/m^3\)) can only be made via site-specific calibration against mass measurements from collocated gravimetric sampling devices. (EPA (1996), Vol. I, 1-6). The correlation between the different particle measures may have seasonal, meteorological, and geographical variations, and the fact that various particle mass measures are employed in different studies therefore complicates using any particular particle measure as indicator of airborne particulates. Some measurement error is necessarily induced by using common converters.

Each study in the meta-analysis supplied mean values of daily data over the study period (often from several monitoring sites) for either TSP, BS, or PM. TSP and PM\textsubscript{13} were converted to PM\textsubscript{10} using the factors of 0.55 and 0.77, respectively, and black smoke was considered equal to PM\textsubscript{10}. Note that this implied dividing the estimated coefficients in studies using the TSP and PM\textsubscript{13} measures by 0.55 and 0.77, respectively, in order to convert these into being PM\textsubscript{10} or BS effects. When converting TSP to PM\textsubscript{10} we relied on the estimate of EPA\textsuperscript{101} which suggested that PM\textsubscript{10} is between 0.5 and 0.6. of TSP. We chose the mean of

\textsuperscript{101} See EPA (1982).
0.55 as our conversion factor. As for BS, data from co-located BS and TSP monitors suggest an average ratio of BS/TSP of 0.55, and it is therefore assumed BS is roughly equivalent to PM$_{10}$. The conversion factor for PM$_{10}$ to PM$_{10}$ was simply obtained by dividing 10 by 13. A few studies used both BS and TSP as particle measures, and in these cases we chose the TSP measure, a gravimetric measure and therefore more straightforward to convert to PM$_{10}$. Particles in ambient air are usually divided into two groups according to size: fine (diameter less than 2.5 μm) and coarse (diameter larger than 2.5 μm). The two size fractions tend to have different origins, composition, and health effects and this makes conversions from fine particle measures to coarse problematic. PM$_{2.5}$ and COH are essentially fine particle measures, and studies using these measures have been excluded from the present analysis.

The proportion of population over 65 (OVER65) was used as a measure of the segment of the population that empirically has been found to be most at risk from the acute effects of air pollution. These data were obtained on a country-level from the World Bank (SIMA). The SIMA data-base provided yearly observations on the percentage of the population over 65 years of age for all the study countries and all the required years. The OVER65-measure used in our regression analysis is hence the average for the relevant study period. Studies carried out in the same country may therefore have different OVER65-measures because they were carried out in different time periods. Three cautionary remarks are in order. First the studies are carried out in specific geographical entities within a country that do not necessarily have the same age distribution in their population as the country overall and this may therefore introduce some degree of measurement error into our regression analysis. Second, the impacts of air pollution on deaths by age group may be very different in low-income than in high-income countries. Cropper et al (1997) found that in Delhi peak effects occurred in the 15 to 44 age group, whereas in the US peak effects occur among people 65 and older. Finally, certain studies have also found that young children may be more susceptible than the average population to high levels of air pollution. A large proportion of people over 65 will tend to be negatively correlated with the proportion of young children, and this may thus bias the OVER65 variable downwards.

GNP per capita at purchasing power parity (PPP) is used as a measure of average

---

102 See Cummings and Waller (1967).
103 SIMA is the World Bank’s internal database system containing more than 40 databases from the Bank and other international institutions.
income in the regression analysis. PPP GNP is gross national product converted to international dollars using purchasing power parity rates. An international dollar has the same purchasing power over GNP as a U.S. dollar has in the United States (i.e. the same amounts of goods and services can be purchased in the domestic market as a U.S. dollar can in the United States). For the restricted meta-analysis we used PPP GNP for 1995 taken from the World Development Report 1997. For the full meta-analysis, updated PPP GNP were obtained from the World Bank (SIMA). The SIMA data-base provided yearly observations for most of the study countries from 1975 onwards. The income measure used in our regression analysis is hence an average for the relevant study period. Main weakness of the measure is the fact that the income level in the location where the study was carried out may differ significantly from the overall income level of the country.

The GINI-coefficient was used to measure inequality in the income distribution of a country. The Gini-coefficient measures the extent to which the distribution of income among individuals or households within an economy deviates from an equal distribution. A Lorenz curve plots the cumulative percentages of total income received against the cumulative number of recipients, starting with the poorest household. The Gini-coefficient measures the area between the Lorenz curve and the line of absolute equality, expressed as a percentage of the maximum area under the line. Hence, a Gini-coefficient of zero represents perfect equality, and an index of 100 implies perfect inequality (World Development Indicators 2000). The coefficients used in the MG meta-analysis were found in the World Development Report 1998/99. For the full meta-analysis, updated gini-coefficients were obtained from the World Bank (SIMA). In the SIMA data-base data on the income distribution within countries were available for all of the countries included in the meta-analysis. It is important to note, however, that the number of observations over time is very limited for most countries, and furthermore that national data differ greatly in terms of how data are collected and expressed (e.g. are the coefficients calculated for income or consumption, gross income or taxable income, household income or individual income?). Furthermore, the income distribution of the cities in the meta-studies are not necessarily the same as the overall income distribution of their respective countries. The GINI-

104 Purchasing power parity conversion factor is the number of units of a country's currency required to buy the same amounts of goods and services in the domestic market as U.S. dollar would buy in the United States. Purchasing power parity conversion factors are estimates by World Bank staff based on data collected by the International Comparison Programme (World Development Indicators 2000).

105 GINI for Greece was not supplied and had therefore to be estimated based on a table from the Greek National Statistical Service (1994), provided by George Korres, Assistant Professor, Athens University. The coefficient for South-Korea was not available, but was found in a paper by Renwei and Shi (1997).
coefficients will therefore most probably measure income distribution with some degree of error.

The interaction term between the GINI-coefficient and GNP per capita, DIST, will reveal whether the effect of the distribution of income on the slope of the dose-response function differs between low- and high-income countries.

Several measures of education were considered; enrolment ratios (education participation), expected years of schooling and illiteracy rates (education outcomes), as well as indicators for education efficiency. Out of these indicators only data on net and gross enrolment ratios were available for a large number of countries (and all of the countries included in the analysis). The gross enrolment ratio is the ratio of total enrolment, regardless of age, to the population of the age group corresponding to the relevant level of education, whereas the net enrolment ratio is the ratio of the number of children of official school age actually enrolled in school to the population of the corresponding official school age. Because the gross enrolment ratio necessarily also includes repeaters, a high ratio does not necessarily indicate a successful education system. For this reason we have chosen net enrolment as the preferred education/knowledge indicator. A drawback of the latter indicator is that children who start school at an age earlier or later than the official school age will not be included in this ratio. More generally, enrolment does not reflect actual attendance, and there may be reasons for overstating enrolments if for example teacher pay is related to student enrolment. Two net enrolment ratios were available; one for primary and one for secondary education. Net enrolment in secondary education was chosen as our education indicator (EDUC) because the majority of the countries in our sample had a net primary enrolment ratio of 100 percent, rendering the latter indicator powerless as a moderator variable. Observations on net secondary enrolment ratios were available for all the countries in the analysis back to 1980. The data for net secondary enrolment ratio was once again obtained from SIMA, and were available from 1980 onwards. They were averaged over the relevant study period. A measurement error may have been introduced due to the fact that enrolment ratios locally may differ from country-level ratios.

Two principal approaches are used to provide summary measures of population health. Disability-Adjusted Life Expectancy (DALE) summarises the expected number of years to be lived in the equivalent of 'full health', i.e. adjusted to take account of time lived with a disability or illness. Disability-Adjusted Life Years (DALYs), on the other
hand, are a gap measure; they measure the gap between a population's actual health and some defined goal (a long life free of illness and disability). The relationship between life-expectancy at birth (LEAB), DALE, and DALYs can easily be shown with the help of a graph depicting survival curves (figure 6.1). The survivorship curve (bold line in figure 6.1) indicates, for each age along the x-axis, the proportion of an initial birth cohort that will remain alive at that age. Life expectancy at birth is equal to the total area under the survivorship curve (i.e. it equals areas A+B). Area A is time lived in full health, whereas area B is time lived in a health state that is less than full. Disability-adjusted life expectancy weighs the time spent in B by the severity of the health states that B represents before adding it to the area below the full-health-survivorship curve (i.e. area A). Finally, disability adjusted life years quantify the difference between the actual health of a population and some stated goal for population health (in figure 6.1 the health goal is to live in ideal health until the death-day). DALYs weigh the time spent in B by the severity of the health states that B represents before adding it to the area above the full-health-survivorship curve, i.e. area C. (Mathers et al. (2000)).

DALE is estimated using information on the fraction of the population surviving to each age (calculated from birth and death rates), the prevalence of each type of disability at each age, and the weight assigned to each type of disability. Survival at each age is adjusted downward by the sum of all the disability effects, each of which is the product of a weight and the complement of a prevalence (the share of the population not suffering that disability). The adjusted survival shares are then divided by the initial population to give the average number of equivalent healthy life years that a new-born can expect. If we enumerate health states, $S$, using a discrete index $h$, DALE can be calculated as follows:

$$ DALE_x = \sum_h w_h(u) \times S_h(u) du $$

where $w_h$ is weight, $u$ represents age, and the integral is over ages from $x$ onwards ($L$ represents the end of the life-time).

The DALE estimate for the population of each country was found in the World Health Report, Annex Table 5, of the World Health Organisation. As this is a relatively newly developed health indicator, estimates were available for 1999 only. Although this is an indicator that may not be changing rapidly, it will nevertheless be an unprecise measure of baseline health, especially in the older studies.
Although an individual with low health levels is more likely on average to die relatively early compared to an individual with higher health levels, life-expectancy at birth (LEAB) is an inaccurate measure of population health since it does not take illness and disability into account. The advantage of this measure is that it was available in SIMA, and has been calculated for the countries in our sample with irregular intervals since the 1970’s. LEAB therefore offers the possibility, although imperfect, of adjusting our health measure to reflect the period in which a particular study was carried out.

Other health indicators are either focusing on specific population groups (e.g. infant mortality), specific adverse health occurrences (e.g. per cent of population with HIV), or are at most indirect measures of population health by measuring expenditure on health or health facilities per capita (e.g. amount of inhabitants per hospital or per doctor), and are therefore not interesting for the present purposes.
Finally, for the purpose of testing the sensitivity of some of our results we wanted to include a measure of the countries' health services. We rejected the use of health expenditure data as a measure of availability and quality of health services, on the ground of being a measure of input that would "reward" inefficient health service systems. From the health service indicators and health utilisation indicators supplied in SIMA, only the former (physicians and hospital beds per 1000 people) were available for all of the countries in our sample. The number of physicians per 1000 people was chosen as measure of the health service in a country. Data were available in SIMA, and have been calculated for the countries in our sample with irregular intervals since the 1970's, hence approximated averages could be calculated for the study periods. The main weakness of this measure is that it does not reveal anything about the distribution of these physicians in various regions or income-classes. In addition, some countries incorrectly included retired physicians or those working outside the health sector. The data can be found in a table in appendix 6.D.
6.4.3 Methodology

In this section we will briefly compare two alternative regression methods, derive the log likelihood function for the mixed effect Empirical Bayes model, as well as describe the tests for homogeneity and for outliers.

In order to obtain the coefficients of the moderator variables two alternative regression methods will be described and briefly compared. In Variance-Weighted Least Squares regressions (VWLS), the concentration-response functions are weighted according to the statistical precision of the studies using the inverse of the variance of each study. This is the method used by Maddison and Gaarder (2000) and the ensuing weights for their sample are given in table 6.1. VWLS differs from Ordinary Least Square (OLS) in that homogeneity of variance is not assumed – the conditional variance of the dependent variable is estimated prior to the regression. VWLS treats the estimated variance as if it were the true variance when it computes standard errors. This method implicitly assumes that all the variance among the study effects other than sampling variance can be explained as a function of known study characteristics (i.e. there is no unexplained between-study variability). We consider this an unrealistic assumption, and note that when available knowledge is insufficient to account for the between-study variation, the model is misspecified. The Empirical Bayes method offers a way of dealing with the insufficiency of knowledge, in particular; it allows us to model the variation among the effect sizes as a function of study characteristics plus error. Empirical Bayes is therefore the main method used in this paper.

According to Raudenbush and Bryk (1985), the Empirical Bayes meta-analysis can be considered a special case of a two-stage hierarchical linear model. The first stage consists of estimating a within-study model separately for each study, and at the second stage a between-study model explains variation in the within-unit parameters as a function of differences between units. This distribution of the true effect size consists of a vector of known constants representing differences between the studies, a vector of between-study parameters, and a random error term, and it is referred to as the prior distribution of the true effect size. Empirical Bayes methods provide a general strategy for estimation when many parameters must be estimated and the parameters themselves constitute realisations from a prior probability distribution.

Estimates can differ partly due to the fact that the studies use different samples of the total population and partly due to the differing conditions under which the research takes
place. Fixed effects models assume the existence of a common effect size in all the studies, whereas random effects models assume a different real effect in each study. In the latter case, combining effect sizes from empirical studies means assessing the average size of the real effect. The common or average effect can be found by calculating the variance weighted average of the effect sizes found, and will be called $\beta w$. In order to choose whether the fixed or the random effects model is the most appropriate, we can perform a homogeneity test using Cochran’s Q-statistic defined as:

$$Q = \sum_{i=1}^{k} \left( \frac{\beta_i - \beta_w}{v_i} \right)^2$$  \hspace{1cm} (6.9)

where $v_i$ is the variance of the reported effect from study $i$, $\beta_i$. If the sample size is large in each study, $Q$ asymptotically has a $\chi^2$-distribution with $k-1$ degrees of freedom. The hypothesis of homogeneity will be rejected if the value of $Q$ is large.

If we reject the hypothesis of equal real effect sizes, the next question is then whether we can find moderator variables that explain the variations between the empirically estimated effect sizes. If a linear combination of variables fully explains the variations in the real effect sizes, then the effect size is fixed and not random (although the real effect sizes are different in each study). This is, however, a rare case. In most cases it is more realistic to use a model that takes into account the imperfections of the explanatory model.

Let us briefly recapitulate the main equations for the mixed effect model already presented in section 6.2. We assumed that the estimated effect size $d_i$ of study $i$ is a function of known study characteristics $W_i$, random errors $u_i$ (inter-study variability) and errors of estimate $e_i$ (intra-study variability):

$$d_i = W_i \gamma + u_i + e_i$$

Assuming that the error terms are independent, the marginal distribution of $d_i$ is:

$$d_i \sim N(W_i \gamma, v_i + \tau^2)$$

Raudenbush and Bryk (1985) use maximum likelihood techniques to derive empirical Bayes estimates “because these techniques are more widely understood than Bayesian methods”. If we assume that the estimate of $v_i$ from each study is approximately equivalent to its true value, we can find the likelihood of the data as a function of $\tau^2$ alone, and thereby find the likelihood estimate of $\tau^2$.

Following Raudenbush and Bryk, $\tau^2$ is determined by maximum likelihood method, where the log of the likelihood is proportional to:
Furthermore, $\gamma^*$ is the maximum likelihood estimate for the vector of derived coefficients, and is given by the following expression:

$$\gamma^* = \left( \sum \lambda_i W_i \right) \sum \lambda_i W_i \beta_i \text{ where } \lambda_i = \frac{\tau^2}{(v_i + \tau^2)}$$

The mathematical derivation of these results is presented in appendix 6.B.

We developed a new programme in STATA (version 6) in order to maximise the above likelihood function, which can be found in appendix 6.C.

There are three key issues in identifying model sensitivity to individual observations, and these are known as residuals, leverage, and influence. The residuals reveal the distance between the value of the $i$th dependent variable, $Y_i$, and the fitted value, $\hat{Y}_i$, and an outlier is identified by a large residual. The leverage, on the other hand, reveals the distance between the value of the independent variable for the $i$th observation, $X_i$, and the mean of all the $X$ values, $\bar{X}$. Having a large leverage can hence also identify an outlier. However, points with large residuals may, but need not, have a large effect on the results, and points with small residuals may still have a large effect, and similarly for the leverage. ‘Influential’ is therefore defined with respect to an index that is affected by the size of the residuals and the size of the leverage. Two outlier tests were performed on our sample. The first test, suggested by Belsley, Kuh, and Welsch (1980), requires that DFITS values greater than $2\sqrt{k/n}$ are subjected to further investigation. The DFITs can be written as follows:

$$DFITS_i = r_i \sqrt{\frac{h_i}{1-h_i}}$$

where $r_i$ are the residuals, $h_i$ is the $i$th leverage, $k$ is the number of explanatory variables (including the constant), and $n$ is the number of observations. DFITS is an attempt to summarise the information in the leverage versus residual-squared-plot into a single statistic. The second test, known as Welsch’s Distance $W_i$, is defined as follows:

$$W_i = DFITS_i \sqrt{\frac{n-1}{1-h_i}}$$

The cutoff for Welsch Distance is $3\sqrt{k}$. 

$$- \sum \log(v_i + \tau^2) - \log\left| \sum (v_i + \tau^2)^{-1} W_i \right| - \sum (v_i + \tau^2)^{-1} (d_i - W_i \gamma^*)^2 \quad (6.10)$$
6.5 Results

6.5.1 The MG Sample

The regression coefficients arising from regressing mortality on PM$_{10}$ in each study were in the study by Maddison and Gaarder (2001) explained by reference to a constant term (CONSTANT), a variable representing the percentage of population over the age of 65 years (OVER65), a variable that represents GNP per capita at PPP (GNP), an interaction term composed of the product of GNP per capita at PPP and the GINI coefficient (DIST), and finally a variable representing average pollution concentrations in each location (POLL).

The original results were first reproduced and are exhibited in table 6.1 (equation 1a). Previous errors in the data set were corrected and the revised regression results are given in equation 2a. The moderator variables were next updated taking the years of the study into account and the updated regression results appear in equation 3a. In equations 1b, 2b, and 3b the same regressions were carried out, however, the Sao Paolo study was excluded from the sample. As argued in section 6.4, it is not correct to include a study looking exclusively at individuals over 65 years of age.

By comparing the first and second regression results in table 6.1 we see that the main effect of the correction in the data set was a decrease in the significance and (negative) effect of GNP on the dose-response coefficient, $\beta$, as well as a decrease in the effect of DIST and POLL. GNP and POLL both remain insignificant, whereas DIST remains significant. When we in addition include the updated moderator variables in regression equation 3a the effects of GNP and DIST increase and both are highly significant, whereas there is a decrease in the estimated effect of OVER65 and POLL and they both become insignificant.

When performing the same regressions on the corrected sample, i.e. excluding Sao Paolo, we observe that the main difference occurs for the OVER65 variable, which becomes insignificant in all three regressions. As compared to the larger sample, excluding Sao Paolo does not alter the estimated effect of POLL. When looking at the regression equation that uses the updated data (i.e. equation 3), we observe that the effects of GNP and DIST are relatively unaffected by the exclusion of Sao Paolo, although there

---

$^{106}$ The null-hypothesis of no effect cannot be rejected at the 20 per cent level.

$^{107}$ The null-hypothesis of no effect is rejected at the 5 per cent level.
is a slight decrease in significance. However, both variables remain significant at the 10% level.

Including the additional moderator variables DALE and EDUC in our regression model (equation 4) increases the positive and negative effects of DIST and GNP, respectively, and both remain highly significant. Given the high correlation between DALE, EDUC, and GNP (shown in table 6.3), the income-related variables were most probably proxying for the missing health and education variables, who appear to have contrasting effects on the dose-response coefficient in this sample. POLL and OVER65 remain insignificant. DALE takes a negative coefficient and is weakly significant, whereas EDUC has an estimated positive effect but is insignificant. If, instead of using the interaction term DIST, we employ the pure relative income inequality measure GINI in our regression model (equation 5), some interesting insights emerge. The OVER65 variable goes from being insignificant to becoming significant, and the effect increases. There is similarly a sizeable increase in the (negative) effect of DALE, whose estimated coefficient becomes significant. The GINI-coefficient is positive and highly significant and the effect of GNP is still significantly negative. However, the absolute size of the effect of GNP has decreased. (Pollution and the constant are still insignificant). The total negative effect of GNP estimated in equation 5 appears to be smaller than the direct negative effect (i.e. excluding the interaction-effect with GINI) in equation 4 due to income inequality having a larger increasing effect in high income countries. Furthermore, OVER65, DALE, and EDUC are all highly correlated with GNP, as can be observed in table 6.3, and their effects all loose significance when DIST is used instead of GINI. This is probably because GNP is included twice in regression equation 4.

If the models hold, \( Q \) has a chi-square distribution with \( f - k \) degrees of freedom, so the null-hypothesis \( H_0 \) that the model holds is rejected when \( Q \geq \chi^2_{f-k}(\alpha) \). None of the models presented in table 6.1 can be rejected at the 0.05 level of significance. Regression models 4 and 5 are clearly the preferred models, with model 5 performing slightly better than model 4 (higher model chi-square and more unlikely to be rejected).

In order to observe how the moderator variables interact, regressions were performed with and without the various variables. The most interesting insight was obtained when excluding the baseline health variable, DALE, from the regression (equation 6). Basically, excluding DALE decreased the effect and significance of all the

---

108 The null-hypothesis of no effect rejected at the 20 per cent level, but not at the 10 per cent level.
explanatory variables, and only the income inequality measure, GINI, was left significant. As we can observe in table 6.3 below, and which is furthermore intuitively quite clear, baseline health (DALE) is highly positively correlated with the proportion of the population over 65 years of age (OVER65), net enrolment ratio into secondary schooling (EDUC), and income (GNP), and negatively correlated with the level of air pollution (POLL). Hence, the insignificant coefficient estimates for EDUC, GNP, and OVER65 when DALE is excluded from the regression are most likely due to these variables proxying for the missing baseline health variable.

The effect estimates are presented in appendix D, column 8, as the percentage change in daily mortality associated with a 10 μg/m³ increase in PM10 levels. The effect estimates in the MG-sample range from 0.02 per cent to 1.21 per cent. We used Cochran’s Q-statistic to assess homogeneity of the studies and found significant heterogeneity in effect estimates (the value of the test-statistic is 34.34, which exceeds the 99th percentile point of the Chi-square distribution with 11 degrees of freedom. The pooled random effects estimate for this sample is 0.37 per cent (the variance weighted effects estimate is 0.32 per cent, and the unweighted average is 0.39 per cent).

As we found heterogeneity among the studies, it would be more correct to use the mixed effect EB model, presented in section 6.4, rather than the variance weighted least squares model. Interestingly, we found when including different variations of our moderator variables that only adding the GINI-predictor reduced \( \hat{r} \) from 5.07e-08 to 4.58e-20, i.e. essentially to zero. Whenever GINI or its composite with GNP, DIST, were included as moderator variables the unexplained heterogeneity was removed. The coefficients of the different regression equations in table 6.3 remained unaltered when performing the mixed effect EB regressions; however, the standard errors of all of the variables were reduced as compared to the variance weighted least squares method, and their significance increased (see table 6.4). This essentially implies that no unexplained heterogeneity remained after this moderator variable was added to the constant. The results from the preferred method (equations 7, 8 and 9) suggests that the effects of the percentage of the population over 65, OVER65, and of the income inequality measures, GINI and DIST, are significant and of the expected (positive) sign, whereas the income-variable, GNP, is associated significantly with a decrease in the dose-response coefficient. Rather surprisingly, a positive and weakly significant association is obtained between enrolment in secondary education, EDUC, and the dose-response coefficient.
Comparing the previous results from the Maddison and Gaarder study with the present results, two main effects of correcting and updating the data and the sample (i.e. excluding the Sao Paolo study) stand out. There is a decrease in the effect and significance of both the level of ambient pollution and of the percentage of people over 65 years of age, and an increase in the negative effect of per capita income, which becomes significant.

It is essential to note that with a sample of twelve or thirteen observations, which is the case of the analysis reviewed in the current section, and with four to six moderator variables and a constant, we are left with around five or six degrees of freedom. The greater variability due to small sample size requires that we have a larger confidence interval. Small sample size therefore induces wariness regarding the results, and is in fact likely to drive the results in the MG-sample analysis.

In addition to the more general notes of caution that will be made in connection with the generalisation of results from the full sample section, the MG-sample is not only small but is also likely to be biased and unrepresentative. In particular, more than half of the low-income cities in our sample (with low GNP per capita) are Eastern European with an unusually low GINI-coefficient compared to the average for low-income countries. The potential bias is further increased due to the higher weights given to the Polish cities than the other observations - jointly they make up approximately three-quarters of the study (see table 6.2 below). This latter fact leads to an even more serious problem in relation to the present regression analysis. The analysis is focusing on moderator variables using country level characteristics rather than city-level characteristics, which implies that because the four Polish dose-response coefficients are relatively low, any population characteristic for which the Polish values stand out from the rest will receive large significance.

\[109\] For the low-income countries in our restricted sample the average GINI-coefficient is 34.3 (where 100 measures full inequality), whereas the average of all available GINI-coefficients from developing countries is 41.4 (World Development Report, 1998/1999).
Table 6.2: Summary of regression results using VWLS-methodology (MG-sample).

<table>
<thead>
<tr>
<th>Equation Method</th>
<th>1a</th>
<th>2a</th>
<th>3a</th>
<th>1b</th>
<th>2b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dep. var. OBS</td>
<td>VWLS</td>
<td>VWLS</td>
<td>VWLS</td>
<td>VWLS</td>
<td>VWLS</td>
</tr>
<tr>
<td>CONSTANT</td>
<td>-0.1411 (0.1252)</td>
<td>-0.1169 (0.1611)</td>
<td>0.0535 (0.1998)</td>
<td>-0.3648 (0.2862)</td>
<td>-0.1492 (0.3002)</td>
</tr>
<tr>
<td>POLL</td>
<td>0.00210' (0.00116)</td>
<td>0.00174 (0.00131)</td>
<td>0.00091 (0.00141)</td>
<td>0.00248** (0.00124)</td>
<td>0.00180 (0.00138)</td>
</tr>
<tr>
<td>OVER65</td>
<td>0.00799** (0.00285)</td>
<td>0.00792** (0.00288)</td>
<td>0.0021 (0.00332)</td>
<td>0.0384 (0.0351)</td>
<td>0.01248 (0.0360)</td>
</tr>
<tr>
<td>GNP</td>
<td>-0.0000001 (0.0000225)</td>
<td>-0.000000225 (0.0000378)</td>
<td>-0.0000124*** (0.0000395)</td>
<td>-0.00000252 (0.0000451)</td>
<td>-0.00000072 (0.0000451)</td>
</tr>
<tr>
<td>DIST</td>
<td>1.77e-06** (5.84e-07)</td>
<td>1.62e-06** (6.10e-07)</td>
<td>4.85e-06*** (1.09e-06)</td>
<td>2.24e-06** (7.92e-07)</td>
<td>1.70e-06* (8.66e-07)</td>
</tr>
<tr>
<td>DALE</td>
<td>(3.03) (2.66)</td>
<td>(2.66) (4.47)</td>
<td>(2.82) (4.47)</td>
<td>(2.82) (4.47)</td>
<td>(2.82) (4.47)</td>
</tr>
<tr>
<td>EDUC</td>
<td>Z</td>
<td>Z</td>
<td>Z</td>
<td>Z</td>
<td>Z</td>
</tr>
<tr>
<td>Goodness-of-fit chi2</td>
<td>9.68</td>
<td>7.59</td>
<td>5.22</td>
<td>8.92</td>
<td>7.57</td>
</tr>
<tr>
<td>Model chi2</td>
<td>48.53</td>
<td>39.40</td>
<td>41.78</td>
<td>34.44</td>
<td>26.76</td>
</tr>
</tbody>
</table>

Source: Own regression results.

Note: The dose-response coefficient, $\beta$, and its standard error have both been multiplied by 1000 in order to make the numbers more readable.
The standard error is given in parenthesis below the estimated coefficients, and the z-statistic is in italics and parenthesis.
Table 6.1 continued: Summary of regression results using VWLS-methodology (MG-sample).

<table>
<thead>
<tr>
<th>Equation</th>
<th>3b</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
<td>VWLS</td>
<td>VWLS</td>
<td>VWLS</td>
<td>VWLS</td>
</tr>
<tr>
<td>Dep. var.</td>
<td>β</td>
<td>β</td>
<td>β</td>
<td>β</td>
</tr>
<tr>
<td>OBS</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>CONSTANT</td>
<td>0.0346</td>
<td>2.414</td>
<td>2.8583</td>
<td>-1.3418</td>
</tr>
<tr>
<td>se</td>
<td>(0.3287)</td>
<td>(1.908)</td>
<td>(1.9773)</td>
<td>(1.0388)</td>
</tr>
<tr>
<td>z</td>
<td>(0.11)</td>
<td>(1.27)</td>
<td>(1.45)</td>
<td>(1.29)</td>
</tr>
<tr>
<td>POLL</td>
<td>0.00094</td>
<td>-0.800277</td>
<td>-0.00164</td>
<td>0.00183</td>
</tr>
<tr>
<td>se</td>
<td>(0.00147)</td>
<td>(0.00215)</td>
<td>(0.00241)</td>
<td>(0.00197)</td>
</tr>
<tr>
<td>z</td>
<td>(0.64)</td>
<td>(-0.13)</td>
<td>(-0.68)</td>
<td>(0.93)</td>
</tr>
<tr>
<td>OVER65</td>
<td>0.00533</td>
<td>0.0775</td>
<td>0.2691*</td>
<td>0.0415</td>
</tr>
<tr>
<td>se</td>
<td>(0.0447)</td>
<td>(0.0621)</td>
<td>(0.1065)</td>
<td>(0.0550)</td>
</tr>
<tr>
<td>z</td>
<td>(0.12)</td>
<td>(1.25)</td>
<td>(2.53)</td>
<td>(0.76)</td>
</tr>
<tr>
<td>GNP</td>
<td>-0.000127*</td>
<td>-0.000233**</td>
<td>-0.0000914'</td>
<td>-7.64e-07</td>
</tr>
<tr>
<td>se</td>
<td>(0.000055)</td>
<td>(0.000088)</td>
<td>(0.000051)</td>
<td>(0.000036)</td>
</tr>
<tr>
<td>z</td>
<td>(-2.31)</td>
<td>(-2.63)</td>
<td>(-1.80)</td>
<td>(-0.02)</td>
</tr>
<tr>
<td>GINI</td>
<td>0.06445***</td>
<td>0.0243***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>se</td>
<td>0.0191</td>
<td>(0.0103)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>z</td>
<td>(3.37)</td>
<td>(2.36)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIST</td>
<td>4.87e-06***</td>
<td>7.44e-06***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>se</td>
<td>(1.13e-06)</td>
<td>(2.28e-06)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>z</td>
<td>(3.27)</td>
<td>(3.37)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DALE</td>
<td>-0.0706'</td>
<td>-0.1376*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>se</td>
<td>(0.0404)</td>
<td>(0.0551)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>z</td>
<td>(-1.75)</td>
<td>(-2.50)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EDUC</td>
<td>0.0247</td>
<td>0.0357'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>se</td>
<td>(0.0210)</td>
<td>(0.0213)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>z</td>
<td>(1.18)</td>
<td>(1.68)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goodness-of-fit chi2</td>
<td>5.21</td>
<td>2.15</td>
<td>1.45</td>
<td>7.68</td>
</tr>
<tr>
<td>Model chi2</td>
<td>29.13</td>
<td>32.19</td>
<td>32.88</td>
<td>26.65</td>
</tr>
</tbody>
</table>

Source: Own regression results.

Note: The dose-response coefficient, $\beta$, and its standard error have both been multiplied by 1000 in order to make the numbers more readable.
The standard error is given in parenthesis below the estimated coefficients, and the $z$-statistic is in italics and parenthesis.
Table 6.3: Weight given to the studies of the sample with VWLS (variance weighted least squares).\textsuperscript{110}

<table>
<thead>
<tr>
<th>Country</th>
<th>City</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greece</td>
<td>Athens</td>
<td>1.20</td>
</tr>
<tr>
<td>Spain</td>
<td>Barcelona</td>
<td>0.37</td>
</tr>
<tr>
<td>Slovakia</td>
<td>Bratislav</td>
<td>0.11</td>
</tr>
<tr>
<td>Poland</td>
<td>Cracow</td>
<td>1.89</td>
</tr>
<tr>
<td>Poland</td>
<td>Wroclaw</td>
<td>1.41</td>
</tr>
<tr>
<td>Poland</td>
<td>Lodz</td>
<td>2.97</td>
</tr>
<tr>
<td>Poland</td>
<td>Poznan</td>
<td>2.39</td>
</tr>
<tr>
<td>UK</td>
<td>London</td>
<td>0.15</td>
</tr>
<tr>
<td>Germany</td>
<td>Köln</td>
<td>0.17</td>
</tr>
<tr>
<td>France</td>
<td>Lyon</td>
<td>0.05</td>
</tr>
<tr>
<td>India</td>
<td>Delhi</td>
<td>0.44</td>
</tr>
<tr>
<td>Chile</td>
<td>Santiago</td>
<td>0.86</td>
</tr>
</tbody>
</table>

Source: Own calculations.

Table 6.4: Correlations between predictors considered in the restricted meta-analysis (n=12). Values greater than 0.5 are in italics.

<table>
<thead>
<tr>
<th></th>
<th>over65</th>
<th>educ</th>
<th>dale</th>
<th>Leab</th>
<th>gnp</th>
<th>gini</th>
<th>dist</th>
<th>poll</th>
<th>se</th>
</tr>
</thead>
<tbody>
<tr>
<td>over65</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>educ</td>
<td>0.74</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dale</td>
<td>0.81</td>
<td>0.88</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>leab</td>
<td>0.75</td>
<td>0.86</td>
<td>0.99</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gnp</td>
<td>0.93</td>
<td>0.73</td>
<td>0.74</td>
<td>0.66</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gini</td>
<td>-0.22</td>
<td>-0.03</td>
<td>0.15</td>
<td>0.24</td>
<td>-0.11</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dist</td>
<td>0.87</td>
<td>0.71</td>
<td>0.78</td>
<td>0.74</td>
<td>0.95</td>
<td>0.19</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>poll</td>
<td>-0.78</td>
<td>-0.78</td>
<td>-0.79</td>
<td>-0.72</td>
<td>-0.69</td>
<td>0.31</td>
<td>-0.60</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>se</td>
<td>0.40</td>
<td>0.65</td>
<td>0.30</td>
<td>0.27</td>
<td>0.55</td>
<td>-0.13</td>
<td>0.52</td>
<td>-0.36</td>
<td>1</td>
</tr>
</tbody>
</table>

Source: Own regression results.

\textsuperscript{110} The weights were derived by taking the inverse of the variance (which is the squared standard error) and normalise the weights so that they all add up to 12, which is the sample size.
Table 6.5: Summary of regression results using EB-methodology (MG-sample).

<table>
<thead>
<tr>
<th>Method</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dep. var.</td>
<td>β</td>
<td>β</td>
<td>β</td>
</tr>
<tr>
<td>CONSTANT</td>
<td>2.414*</td>
<td>2.858*</td>
<td>6.372**</td>
</tr>
<tr>
<td>se</td>
<td>(1.349)</td>
<td>(1.398)</td>
<td>(2.207)</td>
</tr>
<tr>
<td>z</td>
<td>(1.79)</td>
<td>(2.04)</td>
<td>(2.89)</td>
</tr>
<tr>
<td>POLL</td>
<td>-0.000277</td>
<td>-0.00164</td>
<td>-0.000504</td>
</tr>
<tr>
<td>se</td>
<td>(0.00152)</td>
<td>(0.00171)</td>
<td>(0.00153)</td>
</tr>
<tr>
<td>z</td>
<td>(-0.18)</td>
<td>(-0.96)</td>
<td>(-0.33)</td>
</tr>
<tr>
<td>OVER65</td>
<td>0.0775*</td>
<td>0.269***</td>
<td>0.393***</td>
</tr>
<tr>
<td>se</td>
<td>(0.0439)</td>
<td>(0.0752)</td>
<td>(0.103)</td>
</tr>
<tr>
<td>z</td>
<td>(1.76)</td>
<td>(3.57)</td>
<td>(3.84)</td>
</tr>
<tr>
<td>GNP</td>
<td>-0.000233***</td>
<td>-0.0000914*</td>
<td>-0.000168****</td>
</tr>
<tr>
<td>se</td>
<td>(0.0000624)</td>
<td>(0.0000361)</td>
<td>(0.0000517)</td>
</tr>
<tr>
<td>z</td>
<td>(-3.72)</td>
<td>(-2.53)</td>
<td>(-3.25)</td>
</tr>
<tr>
<td>GINI</td>
<td>0.0645***</td>
<td>0.0827***</td>
<td></td>
</tr>
<tr>
<td>se</td>
<td>(0.0135)</td>
<td>(0.01736)</td>
<td></td>
</tr>
<tr>
<td>z</td>
<td>(4.77)</td>
<td>(4.76)</td>
<td></td>
</tr>
<tr>
<td>DIST</td>
<td>7.44e-06***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>se</td>
<td>(1.61e-06)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>z</td>
<td>(4.62)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LEAB</td>
<td>-0.2103***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>se</td>
<td>(0.0567)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>z</td>
<td>(-3.71)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DALE</td>
<td>-0.0706*</td>
<td>-0.138***</td>
<td></td>
</tr>
<tr>
<td>se</td>
<td>(0.0286)</td>
<td>(0.0390)</td>
<td></td>
</tr>
<tr>
<td>z</td>
<td>(-2.47)</td>
<td>(-3.53)</td>
<td></td>
</tr>
<tr>
<td>EDUC</td>
<td>0.0247*</td>
<td>0.0357*</td>
<td>0.0486**</td>
</tr>
<tr>
<td>se</td>
<td>(0.0148)</td>
<td>(0.0150)</td>
<td>(0.0170)</td>
</tr>
<tr>
<td>z</td>
<td>(1.67)</td>
<td>(2.38)</td>
<td>(2.86)</td>
</tr>
<tr>
<td>Tau2</td>
<td>6.61e-23</td>
<td>2.99e-22</td>
<td>2.20e-15</td>
</tr>
<tr>
<td>se</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>z</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Log like.</td>
<td>-31.49</td>
<td>-32.72</td>
<td>-31.33</td>
</tr>
<tr>
<td>Wald chi2</td>
<td>64.37</td>
<td>65.77</td>
<td>67.05</td>
</tr>
</tbody>
</table>

'P<0.1, *P<0.05, **P<0.01, ***P<0.001

Source: Own regression results.

Note: The dose-response coefficient, β, and its standard error have both been multiplied by 1000 in order to make the numbers more readable. The standard error is given in parenthesis below the estimated coefficients, and the z-statistic is in italics and parenthesis. Tau2 is the inter-study variation.
6.5.2 Full Sample

In our full sample we used the coefficients reported, no matter what lag structure was used and whether additional pollutants were included in the model or not. If several coefficients were reported we used the one favoured by the researcher, and if no preference was mentioned we chose the most significant coefficient. In the 8 studies reporting results both for single and multiple pollutants we used the preferred single pollutant results, since the large majority of studies only reported single pollutant results. The Aphea group decided to search for the best relationship of each pollutant with mortality with a lag of up to 3 days, and the best cumulative effect of several days including the same day and up to 3 previous days, and the EMECAM project in Spain followed the methodology developed by the Aphea group. In practice, this lead to a sample largely consisting of single pollutant models.

We used Cochran's $Q$-statistic on the pooled random effects model to assess homogeneity of the studies, and found significant heterogeneity in effect estimates. The value of the test-statistic is 329, exceeding the 99th percentile point of the Chi-square distribution with 69 degrees of freedom. Applying the two outlier-tests described in section 6.5.3 to the pooled random effects model, we found that no observation in the sample failed the Welsch's Distance test. Three observations (Huelva, Pamplona, and Basel) failed the DFITS cutoff point; however, the heterogeneity in effect estimates remained even after these were excluded from the sample (the value of Cochran's $Q$-statistics was 301).

When we apply the EB model to the full sample of 70 observations, a model without predictors yields an estimated grand mean of the regression coefficients of 0.00060, which is equivalent to a 6.0 percentage (CI: 5.1 - 6.9 per cent) increase in daily mortality associated with a 100-µg/m³ increase in $PM_{10}$ concentrations. This pooled estimate is similar to the ones found by Levy et al. (2000), Borja-Aburto et al. (2000), and Ostro (1993), although slightly lower.

When moderator variables were entered one at a time into the model which included only a constant, the income inequality variable, GINI, had the greatest significance from the range of moderator variables available ($z=3.76$). When this moderator variable was added to the model the $\tau^2$ decreased from 1.83e-07 to 1.61e-07 (for the other moderator variables the following $\tau^2$ values were obtained: DIST $\tau^2=1.68e-07$ ($z=3.27$); GNP $\tau^2=1.77e-07$.
By comparing equations 0a and 2a and 0b and 2b in table 6.5, it is clear that practically the same results are obtained whether we use DALE or LEAB as the baseline health measure. We therefore choose DALE as baseline health indicator from this point onwards in our analysis because it is the theoretically preferred indicator. Our main finding is the positive and significant association of the GINI coefficient and its interaction term with GNP, DIST, with the dose-response coefficient. In particular, we find that a four-point increase in the GINI-coefficient leads to an increase in the mortality rate from a 100-μg/m³ increase in PM₁₀ of almost 1 percent. When DIST is included in the model (equations 0a and 2a) we furthermore observe that the coefficient on GNP takes a negative sign but is insignificant, whereas when GINI is included instead (equations 0b and 2b) the GNP-coefficient becomes positive and significant at the 0.05 level. To simplify matters we will from here onwards refer to the former model-specification as the DIST-model, and the latter as the GINI-model. If we do not include any variable picking up income distribution in our regression (equation 1), the average level of income, GNP, is found to have a positive and significant effect on the dose-response association. This seems to indicate that the direct effect of GNP on the dose-response coefficient is negligible, whereas its main effect is by interacting with income inequality. In particular, high income-inequality appears to have a larger (increasing) effect on the dose-response coefficient in high-income countries.

The education term is consistently taking a negative coefficient, but its z-statistic indicates that it is slightly below a level that can be termed significant. As for baseline health, measured either by standard or disability adjusted life expectancy, its coefficient varies but it is insignificant, as are both the constant and the level of average ambient pollution. Somewhat surprisingly, the percentage of the population aged 65 or older is consistently negative, however it is only significant in the case where income is included without any measure of distribution present in the regression (equation 1).

Finally, we included the number of physicians per 1000 people, PHYS, as a moderator variable in equations 4a and 4b. The effect of this variable was found to be negative but insignificant, and the only implication of including it in our regression model worth mentioning was a slight decrease in significance for both GINI and DIST. Both, however, remained significant.
Note that our findings would have been entirely different had we used the OLS or VWLS regression methods. This confirms how important it is to understand the assumptions underlying each method and to choose the correct one. See appendix 6.E for the findings applying the main regression equations on the full sample using the methods of OLS and VWLS.
Table 6.6: Summary of regression results using EB-methodology on full sample (70).

<table>
<thead>
<tr>
<th>Equation</th>
<th>Method</th>
<th>Dep. var.</th>
<th>OBS</th>
<th>0a β</th>
<th>0b β</th>
<th>1 β</th>
<th>2a β</th>
<th>2b β</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONST</td>
<td>EB</td>
<td>70</td>
<td>-0.2417</td>
<td>-0.10689</td>
<td>-0.43615</td>
<td>-0.3296</td>
<td>0.3686</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(1.3268)</td>
<td>(1.3091)</td>
<td>(1.329)</td>
<td>(1.8351)</td>
<td>(1.8300)</td>
<td></td>
</tr>
<tr>
<td>POLL</td>
<td>EB</td>
<td>70</td>
<td>-0.00084</td>
<td>-0.00123</td>
<td>-0.00068</td>
<td>-0.000783</td>
<td>-0.00134</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.00142)</td>
<td>(0.00141)</td>
<td>(0.00142)</td>
<td>(0.00145)</td>
<td>(0.00144)</td>
<td></td>
</tr>
<tr>
<td>OVER65</td>
<td>EB</td>
<td>70</td>
<td>-0.03609</td>
<td>-0.01917</td>
<td>-0.06204*</td>
<td>-0.02909</td>
<td>-0.01271</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.02745)</td>
<td>(0.02908)</td>
<td>(0.02465)</td>
<td>(0.02604)</td>
<td>(0.02727)</td>
<td></td>
</tr>
<tr>
<td>GNP</td>
<td>EB</td>
<td>70</td>
<td>-0.000024</td>
<td>0.000028*</td>
<td>0.000041**</td>
<td>-0.000027</td>
<td>0.000028*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.000033)</td>
<td>(0.000014)</td>
<td>(0.000013)</td>
<td>(0.000033)</td>
<td>(0.000014)</td>
<td></td>
</tr>
<tr>
<td>GINI</td>
<td>EB</td>
<td>70</td>
<td>0.02265**</td>
<td>0.02265**</td>
<td>0.02265**</td>
<td>0.02265**</td>
<td>0.02265**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0085)</td>
<td>(0.0085)</td>
<td>(0.0085)</td>
<td>(0.0085)</td>
<td>(0.0085)</td>
<td></td>
</tr>
<tr>
<td>DIST</td>
<td>EB</td>
<td>(2.67)</td>
<td>1.52e-06*</td>
<td>1.52e-06*</td>
<td>1.52e-06*</td>
<td>1.52e-06*</td>
<td>1.52e-06*</td>
<td></td>
</tr>
<tr>
<td>LEAB</td>
<td>EB</td>
<td>(2.15)</td>
<td>0.02259</td>
<td>0.02844</td>
<td>0.02844</td>
<td>0.02844</td>
<td>0.02844</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.84)</td>
<td>(0.84)</td>
<td>(0.84)</td>
<td>(0.84)</td>
<td>(0.84)</td>
<td></td>
</tr>
<tr>
<td>DALE</td>
<td>EB</td>
<td>(1.32)</td>
<td>0.0248</td>
<td>0.0248</td>
<td>0.0248</td>
<td>0.0248</td>
<td>0.0248</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.22)</td>
<td>(0.22)</td>
<td>(0.22)</td>
<td>(0.22)</td>
<td>(0.22)</td>
<td></td>
</tr>
<tr>
<td>PHYS</td>
<td>EB</td>
<td></td>
<td>-0.00958</td>
<td>-0.00627</td>
<td>-0.00891</td>
<td>-0.00855</td>
<td>-0.00581</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.00586)</td>
<td>(0.00584)</td>
<td>(0.00587)</td>
<td>(0.00574)</td>
<td>(0.00566)</td>
<td></td>
</tr>
<tr>
<td>EDUC</td>
<td>EB</td>
<td></td>
<td>-0.177***</td>
<td>0.170***</td>
<td>0.179***</td>
<td>0.178***</td>
<td>0.170***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(4.80)</td>
<td>(4.76)</td>
<td>(4.80)</td>
<td>(4.82)</td>
<td>(4.76)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(4.80)</td>
<td>(4.76)</td>
<td>(4.80)</td>
<td>(4.82)</td>
<td>(4.76)</td>
<td></td>
</tr>
<tr>
<td>Tau2</td>
<td>EB</td>
<td></td>
<td>-84.39</td>
<td>-64.34</td>
<td>-59.04</td>
<td>-84.27</td>
<td>-63.94</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(19.27)</td>
<td>(22.20)</td>
<td>(14.69)</td>
<td>(18.61)</td>
<td>(22.19)</td>
<td></td>
</tr>
<tr>
<td>Log like.</td>
<td>EB</td>
<td></td>
<td>-84.39</td>
<td>-64.34</td>
<td>-59.04</td>
<td>-84.27</td>
<td>-63.94</td>
<td></td>
</tr>
<tr>
<td>Wald chi2</td>
<td>EB</td>
<td></td>
<td>(19.27)</td>
<td>(22.20)</td>
<td>(14.69)</td>
<td>(18.61)</td>
<td>(22.19)</td>
<td></td>
</tr>
</tbody>
</table>

Source: Own regression results.

Note: The dose-response coefficient, $\beta$, and its standard error have both been multiplied by 1000 in order to make the numbers more readable. The standard error is given in parenthesis below the estimated coefficients, and the $z$-statistic is in italics and parenthesis. Tau2 is the inter-study variation.
Table 6.5 continued: Summary of regression results using EB-methodology (for full sample and sample excluding outlier (69)).

<table>
<thead>
<tr>
<th>Equation</th>
<th>Method</th>
<th>3a</th>
<th>3b</th>
<th>4a</th>
<th>4b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dep. var.</td>
<td>OBS</td>
<td>EB</td>
<td>EB</td>
<td>EB</td>
<td>EB</td>
</tr>
<tr>
<td>METHOD</td>
<td></td>
<td>β</td>
<td>β</td>
<td>β</td>
<td>β</td>
</tr>
<tr>
<td>CONST</td>
<td>69</td>
<td>-0.2846</td>
<td>-0.12835</td>
<td>-0.44796</td>
<td>-0.2047</td>
</tr>
<tr>
<td>se</td>
<td>(1.2352)</td>
<td>(1.2173)</td>
<td>(1.4429)</td>
<td>(1.4295)</td>
<td></td>
</tr>
<tr>
<td>z</td>
<td>(-0.23)</td>
<td>(-0.11)</td>
<td>(-0.31)</td>
<td>(-0.14)</td>
<td></td>
</tr>
<tr>
<td>POLL</td>
<td>69</td>
<td>-0.000409</td>
<td>-0.00084</td>
<td>-0.00646</td>
<td>-0.00114</td>
</tr>
<tr>
<td>se</td>
<td>(0.001323)</td>
<td>(0.00131)</td>
<td>(0.00153)</td>
<td>(0.00154)</td>
<td></td>
</tr>
<tr>
<td>z</td>
<td>(-0.31)</td>
<td>(-0.64)</td>
<td>(-0.42)</td>
<td>(-0.74)</td>
<td></td>
</tr>
<tr>
<td>OVER65</td>
<td>70 (+ phys)</td>
<td>-0.03758</td>
<td>-0.01903</td>
<td>-0.02997</td>
<td>-0.01651</td>
</tr>
<tr>
<td>se</td>
<td>(0.02571)</td>
<td>(0.02724)</td>
<td>(0.03151)</td>
<td>(0.03223)</td>
<td></td>
</tr>
<tr>
<td>z</td>
<td>(-1.46)</td>
<td>(-0.70)</td>
<td>(-0.95)</td>
<td>(-0.51)</td>
<td></td>
</tr>
<tr>
<td>GNP</td>
<td>70 (+ phys)</td>
<td>-0.000032</td>
<td>0.000002</td>
<td>-0.000020</td>
<td>0.000028*</td>
</tr>
<tr>
<td>se</td>
<td>(0.000031)</td>
<td>(0.000013)</td>
<td>(0.000034)</td>
<td>(0.000014)</td>
<td></td>
</tr>
<tr>
<td>z</td>
<td>(-1.04)</td>
<td>(1.54)</td>
<td>(-0.59)</td>
<td>(2.05)</td>
<td></td>
</tr>
<tr>
<td>GINI</td>
<td>70 (+ phys)</td>
<td>0.02356**</td>
<td>0.02213*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>se</td>
<td>(0.00793)</td>
<td>(0.0091)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>z</td>
<td>(2.97)</td>
<td>(2.42)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIST</td>
<td>70 (+ phys)</td>
<td>1.54e-06*</td>
<td>1.42e-06*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>se</td>
<td>(6.70e-07)</td>
<td>(7.68e-07)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>z</td>
<td>(2.30)</td>
<td>(1.85)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LEAB</td>
<td>70 (+ phys)</td>
<td>0.01928</td>
<td>-0.00144</td>
<td>0.02756</td>
<td>0.00685</td>
</tr>
<tr>
<td>se</td>
<td>(0.02034)</td>
<td>(0.0216)</td>
<td>(0.02315)</td>
<td>(0.02527)</td>
<td></td>
</tr>
<tr>
<td>z</td>
<td>(0.95)</td>
<td>(-0.07)</td>
<td>(1.19)</td>
<td>(0.27)</td>
<td></td>
</tr>
<tr>
<td>PHYS</td>
<td>70 (+ phys)</td>
<td>-0.03089</td>
<td>-0.03089</td>
<td>-0.01437</td>
<td></td>
</tr>
<tr>
<td>se</td>
<td>(0.07765)</td>
<td>(0.07688)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>z</td>
<td>(-0.40)</td>
<td>(-0.19)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EDUC</td>
<td>70 (+ phys)</td>
<td>-0.00385</td>
<td>-0.00046</td>
<td>-0.00947</td>
<td>-0.00634</td>
</tr>
<tr>
<td>se</td>
<td>(0.00560)</td>
<td>(0.00557)</td>
<td>(0.00594)</td>
<td>(0.00590)</td>
<td></td>
</tr>
<tr>
<td>z</td>
<td>(-0.69)</td>
<td>(-1.03)</td>
<td>(-1.59)</td>
<td>(-1.07)</td>
<td></td>
</tr>
<tr>
<td>Tau2</td>
<td>70 (+ phys)</td>
<td>0.148***</td>
<td>0.141***</td>
<td>0.182***</td>
<td>0.176***</td>
</tr>
<tr>
<td>se</td>
<td>(4.80)</td>
<td>(4.77)</td>
<td>(4.80)</td>
<td>(4.75)</td>
<td></td>
</tr>
<tr>
<td>z</td>
<td>(4.80)</td>
<td>(4.77)</td>
<td>(4.80)</td>
<td>(4.75)</td>
<td></td>
</tr>
<tr>
<td>Log like.</td>
<td></td>
<td>-73.96</td>
<td>-53.43</td>
<td>-88.74</td>
<td>-68.77</td>
</tr>
<tr>
<td>Wald chi2</td>
<td></td>
<td>17.79</td>
<td>21.73</td>
<td>19.10</td>
<td>21.84</td>
</tr>
</tbody>
</table>

Source: Own regression results.

Note: The dose-response coefficient, β, and its standard error have both been multiplied by 1000 in order to make the numbers more readable. The standard error is given in parenthesis below the estimated coefficients, and the z-statistic is in italics and parenthesis. Tau2 is the inter-study variation.
Table 6.7: Correlations between predictors considered in the full meta-analysis (n=70). Values greater than 0.5 are in italics.

<table>
<thead>
<tr>
<th></th>
<th>over65</th>
<th>educ</th>
<th>Dale</th>
<th>Leab</th>
<th>gnp</th>
<th>gini</th>
<th>dist</th>
<th>poll</th>
<th>phys</th>
<th>se</th>
</tr>
</thead>
<tbody>
<tr>
<td>over65</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>educ</td>
<td>0.60</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dale</td>
<td>0.82</td>
<td>0.65</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>leab</td>
<td>0.80</td>
<td>0.63</td>
<td>0.94</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gnp</td>
<td>0.62</td>
<td>0.66</td>
<td>0.59</td>
<td>0.67</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gini</td>
<td>-0.40</td>
<td>-0.21</td>
<td>-0.17</td>
<td>-0.12</td>
<td>-0.01</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dist</td>
<td>0.45</td>
<td>0.59</td>
<td>0.46</td>
<td>0.55</td>
<td>0.94</td>
<td>0.28</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>poll</td>
<td>-0.53</td>
<td>-0.63</td>
<td>-0.58</td>
<td>-0.63</td>
<td>-0.61</td>
<td>0.16</td>
<td>-0.54</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>phys</td>
<td>0.77</td>
<td>0.41</td>
<td>0.67</td>
<td>0.66</td>
<td>0.32</td>
<td>-0.46</td>
<td>0.14</td>
<td>-0.22</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>se</td>
<td>0.30</td>
<td>0.21</td>
<td>0.27</td>
<td>0.29</td>
<td>0.20</td>
<td>-0.04</td>
<td>0.17</td>
<td>-0.26</td>
<td>0.34</td>
<td>1</td>
</tr>
</tbody>
</table>

Source: Own calculations.
6.5.3 Sensitivity Analysis

There are several reasons why one should be very cautious when generalising the results. First, the full sample is still relatively small, and single observations may affect the results unduly. Second, the conversion factors used between different measures of air pollution are approximations and may have affected the outcome. Also, there is a possible difference in the ‘positive results’-bias between various countries. Furthermore, both the mortality measures and the lag structure considered and/or reported vary between the studies and may have induced errors. Finally, the fact that some of the countries, as well as cities, in the study enter with multiple observations whereas others have only one may be problematic and requires investigation. Given these drawbacks of our sample, a sensitivity analysis is necessary before drawing conclusions about potential causal predictors of the PM$_{10}$-mortality relationship. Special attention is paid below to outliers, mortality measure, negative results, lag structure, pollution measurement, and the number of observations per location.

First, we undertook a search for outliers, i.e. subsets of the observations which, if deleted, would change the results markedly, and tested the sensitivity of our results to exclusion of these observations. When performing the DFITS-test (described in section 6.5.3) on our sample only Basel was singled out. From table 6.6 (equations 3a and 3b) we observe two main changes in estimated effects when excluding Basel from the sample. First, there is an increase in effect and significance of GINI and DIST in equations 3a and 3b, respectively. Second, a decrease in the positive effect of GNP occurs in equation 3b, and an increase in the negative effect of GNP occurs in equation 3a. In both cases GNP is insignificant. No observation in the sample failed Welsch’s Distance test.

Out of the sample of 70 pollution-mortality studies, 9 were not excluding mortality due to external causes from their mortality measure. If externally caused mortality is independent of air pollution, then including these studies in the full sample should not bias the results – they would at most bring more noise into the results. However, to our knowledge no study specifically investigates the association between accidents and homicides etc. and air pollution.

When excluding studies measuring all-cause mortality, the central effect estimate is 5.9 per cent (CI: 5.0 – 6.8 per cent), whereas only looking at all-cause mortality studies gives a mean of 6.5 per cent (CI: 0.13 – 1.18 per cent). Next, we tested whether the
inclusion of studies measuring all-cause mortality induced estimation errors using the two main moderator effect models. In table 6.7 equation 5a we see that the main significant effect of this exclusion was to increase the negative effect of GNP on the dose-response coefficient – this coefficient becomes weakly significant at the 0.1 level – and increase the size and significance of the interaction term, DIST. As for the other terms, their coefficients were similar in magnitude and direction to the optimum model and remain insignificant. When DIST is replaced by GINI (equation 5b), we observe that the increasing effect of GNP on the dose-response coefficient found in the full sample regression decreased and became insignificant through the exclusion of the 9 studies. The increasing effect of GINI on the particle-mortality association, on the other hand, increased further. The remaining coefficients remained insignificant, as compared to the full-sample model, and except for CONSTANT and DALE the coefficients of the additional terms had the same sign.

Out of the pollution-mortality studies, 6 reported negative results, and 4 out of these originated from the EMECAM program studies (i.e. Spain). If all confounding variables were controlled for, the model correctly specified, and the data measured without error, it seems unlikely that we would find negative estimates of the dose-response coefficient, i.e. indicating that air pollution decreases mortality. Hence, we could view the studies reporting negative coefficients as noise in our meta-analysis and therefore leave these out of the analysis.111

The main objection against exclusion of negative results has to do with researcher-induced sample-selection bias. The negative studies may to some extent counteract the noise in positive studies, studies that we would not discard because they have the “right” sign. Furthermore, it is highly likely that if all results had an equal chance of getting published (i.e. without prejudices against negative results) we would in fact have had more negative coefficients in our meta-analysis.

If we exclude the negative coefficients from our sample (hence, sample size 64), the grand mean is 6.7 per cent (CI: 5.9 – 7.5 per cent). The findings from estimating the two

111 An additional, but rather tentative, reason for excluding the negative results is that 4 out of the 6 originated from the EMECAM program studies (i.e. Spain). If all the countries in our analysis had an equal probability of reporting and publishing negative results, then including these studies in the full sample should not affect the relative significance of the moderator variables. However, were this not to be the case, then the ‘positive results’-bias may influence our conclusions. Since Spain and the United States receive the most weight in the meta-analysis due to multiple observations for each country, it is worth noting that the US studies were all published in international and competitive journals, whereas the Spanish studies were both carried out due to a government grant and published in a government-supported journal. The ‘positive results’-bias may therefore well have been larger in the US than in Spain, and for comparability of results one may hence argue that the negative results should be excluded.
main models using the sample consisting only of positive results are similar to those we obtained when we excluded all-cause mortality studies from the sample (equations 6a and 6b).

As mentioned in section 6.4.1, the correlation between the different particle measures may have seasonal, meteorological, and geographical variations, and using fixed and common conversion factors will most likely induce some measurement error.

The central effect estimates when stratifying by pollution measure were 5.0 per cent (CI: 3.4 – 6.7 per cent), 6.6 per cent (CI: 5.1 – 8.2 per cent), and 6.4 per cent (CI: 4.9 – 7.8 per cent) for BS, TSP, and PM10, respectively. We also tested the sensitivity of our findings from the main models to the air particle measurements used, by considering the coefficients derived using TSP, BS, and PM10 separately. A main cautionary remark is called for before entering into the more specific challenges we are faced with when analysing the samples stratified by pollution measurement. The sample sizes for each pollution measurement is relatively small (24 observations for TSP, 33 in the case of PM10, and 16 for the BS measurement), and any results have to be taken with the utmost caution. As most of the moderator variables are country based, rather than city based, and some (GINI and DALE) are available for a certain year only and not in time-series, the regression model developed above is only meaningful when a variety of countries are included in the sample. Furthermore, the larger the amount of studies in the sample originating from the same country, the more likely it is that the variables that do vary from one study to the next proxy for those that do not. Keeping this in mind, and given that 50 per cent of the studies using BS to measure air pollution are from Spain and approximately 50 per cent of those using PM10 have been carried out in the US, we argue that the results from these two samples will not be very meaningful. In the case of the sample of 24 TSP based coefficients, 13 countries are represented and none with more than 5 coefficients. Equations 7a and 7b of table 6.7 give the results of the main model for the TSP-sample. When we compare the results with those of the full sample for the model which includes DIST, we observe a strengthening of the negative effect of GNP on the dose-response coefficient, although it remains insignificant. Furthermore, the size and significance of the interaction term increases. The remaining coefficients once more remain insignificant, however only EDUC which was close to being significant in the full sample retains the direction of its

---

112 The sample sizes for the three air pollution measures estimates were 16, 24, and 33 for BS, TSP and PM10 (including two studies using PM13), respectively. The sample adds up to more than our full sample because 3 of the studies reported estimates in two of the measures.
coefFicient. With GINI replacing DIST (equation 7b), we observe that the positive effect of GNP on the dose-response coefficient found in the full sample regression decreases and becomes insignificant. The positive effect of GINI on the particle-mortality association, on the other hand, increases further. The remaining coefficients once more remain insignificant, and only the coefficients for POLL and EDUC do not change direction. For the sake of completeness the results of the main model for the PM$_{10}$ and BS-sample are given in equations 8 and 9 (table 6.7), respectively.

Next, we investigated the sensitivity of our findings to the choice of lag structure. In particular, we ran our favoured model on samples of the pollution-mortality coefficients that were obtained for average air particle levels of the same day (table 6.8, equations 10a and 10b), previous day (table 6.8, equations 11a and 11b), and of two days previously (lag 2 – equations 12a and 12b) in turn. Before proceeding, however, the same warning has to be made as was made in connection with the pollution measurements analysis on the previous page. The sample sizes for each lag is relatively small (28 observations for lag 1, 23 in the case of lag 1, and 12 for lag 2), and any results have to be taken with the utmost caution, and in particular in the case of the smaller samples is likely to drive the results. Interestingly, the results varied widely. For same day air pollution, the GINI coefficient was once again found to have a strongly significant\textsuperscript{113} positive effect on the dose-response coefficient, whereas GNP took a negative sign but was insignificant. POLL has a significantly negative effect on the relationship. As for baseline health, it is found to have an increasing and significant effect, whereas education has a decreasing and significant effect. For particle pollution lagged one day, none of the moderator terms have significant coefficients. In the case of two-day lagged pollution, however, GINI is once again weakly significant,\textsuperscript{114} with the usual positive sign, and the coefficient for education is significant and this time takes a positive sign. The estimates of central effect were 5.3 per cent (CI: 4.3 – 6.2 per cent), 5.5 per cent (CI: 4.1 – 6.9 per cent), and 3.7 percentage (CI: 1.6 – 5.7 per cent) for lag 0, lag 1, and lag 2, respectively.\textsuperscript{115}

Furthermore, we tested the sensitivity of our findings to the decision to include several studies from the same city, in order to detect the potential effects of double counting. Only 5 cities have been entered twice in the sample, and we tested several combinations

\textsuperscript{113} The null-hypothesis of no effect is rejected at the 0.1 per cent level.

\textsuperscript{114} The null-hypothesis of no effect rejected at the 10 per cent level, but not at the 5 per cent level.

\textsuperscript{115} The sample sizes for lag 0, 1, and 2 were 28, 23, and 12, respectively.
when only one of these studies was considered (sample size was then 65). As expected, the findings did not change significantly from those made for the full sample.

If we had obtained information on the values of the moderator variables at the city level, rather than at the country level, the fact that some countries enter with several observations whereas others only with one would not have been a cause for concern. However, this is not the case – only the average particulate measure was based on the exact study location. Our final sensitivity test therefore consisted of including only one observation for each country. This can be done in three alternative ways. First, the variance weighted averages of both the dependent and independent variables, as well as of the standard error, are calculated for those countries with more than one observation, and the resulting averages can then be entered as observations for those countries. The second procedure is an extension of the first; the only difference is that we take the time period in which the studies were carried out into account, since we have time series data for most of the moderator variables. All studies from one country carried out mainly in the period 1975-85 were averaged, as were those for the periods 1985-95, and 1995-, implying that we had two “observations” each for Spain and Germany, and three for the US, and thus a sample of 25. Finally, we can consider only one (real) study per city, and a number of potential study combinations.

The two main regressions were performed on the simple country-averaged sample of 21 observations (table 6.9, equations 13a and 13b) and on the period sensitive country-averaged sample of 25 observations (table 6.9, equations 14a and 14b). The results from the two samples were similar to each other, and significantly different from our previous results. Once again, however, a cautionary remark is in order when interpreting the results, due to the sample sizes. Both education (EDUC) and the amount of older people in the population (OVER65) were found to have significantly negative effects on the dose-response coefficient, i.e. higher levels of these variables are associated with lower PM_{10}-mortality coefficients. As for the two income distribution terms, they were both insignificant when entered in turn. When entered together with the relative income inequality measure, GINI, the average income-term takes a positive coefficient and is highly significant. As for the income coefficient when the interaction term, DIST, is included in the model, GNP has a significantly positive coefficient in the sample of 21 observations, but is insignificant in the sample of 25. The estimates of the central effect is 4.9 per cent (CI: 3.9 – 5.9 per cent) in the
sample with 21 observations, and 5.0 per cent (CI: 4.1 – 6.0 per cent) in the sample consisting of 25.\(^{116}\)

With over 2 million potential study combinations, a complete combination analysis was not deemed feasible. However, by randomly selecting 24 combinations and performing the two main regressions, we propose that we can gain some insights from which we can generalise.\(^{117}\) Focussing first on the model including the GINI term, we found that GNP entered the regression significantly for 54 per cent of the study combinations, with a consistently positive sign. Education entered with a consistently negative and significant sign for 50 per cent of the study combinations. In addition, DALE, OVER65, and GINI entered in descending order of frequency, although their frequencies were well below 25 per cent. GINI was, in fact, only significant in one study combination and pollution in none. Education, it now entered negatively and significantly in 71 per cent of the study combinations. DALE and OVER65 both again appeared in study combinations at frequencies below 25 per cent.

---

\(^{116}\) When including only one observation for each country, the sample size is 21.

\(^{117}\) Refer to appendix 6.F for the regression results.
Table 6.8: Summary of regression results using EB-methodology (sample excluding all-cause mortality studies (61), sample excluding negative coefficients (64), sample of TSP studies (24), PM10 studies (33), and BS studies (16)).

<table>
<thead>
<tr>
<th>Equation</th>
<th>Method</th>
<th>Dep. var.</th>
<th>OBS</th>
<th>5a</th>
<th>5b</th>
<th>6a</th>
<th>6b</th>
<th>7a</th>
<th>7b</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EB</td>
<td>EB</td>
<td>β</td>
<td>β</td>
<td>β</td>
<td>β</td>
<td>β</td>
<td>β</td>
<td>β</td>
<td>β</td>
<td>β</td>
</tr>
<tr>
<td>CONSTANT</td>
<td>61 (-accid)</td>
<td>61 (-accid)</td>
<td>-0.0101</td>
<td>0.2070</td>
<td>-0.4408</td>
<td>-0.2518</td>
<td>0.7019</td>
<td>0.2659</td>
<td>-1.2173</td>
<td>-4.6017</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.2239)</td>
<td>(1.1940)</td>
<td>(1.1066)</td>
<td>(1.0770)</td>
<td>(2.8333)</td>
<td>(2.9177)</td>
<td>(3.1341)</td>
<td>(3.1584)</td>
<td>(3.1584)</td>
<td>(3.1584)</td>
<td></td>
</tr>
<tr>
<td>POLL</td>
<td>64 (-neg)</td>
<td>64 (-neg)</td>
<td>-0.0090</td>
<td>-0.00149</td>
<td>-0.00114</td>
<td>-0.00162</td>
<td>0.00146</td>
<td>-0.000275</td>
<td>-0.00119</td>
<td>-0.00490'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00131)</td>
<td>(0.00129)</td>
<td>(0.00120)</td>
<td>(0.00118)</td>
<td>(0.00118)</td>
<td>(0.00118)</td>
<td>(0.00118)</td>
<td>(0.00118)</td>
<td>(0.00118)</td>
<td>(0.00118)</td>
<td></td>
</tr>
<tr>
<td>OVER65</td>
<td>24 (TSP)</td>
<td>24 (TSP)</td>
<td>-0.02539</td>
<td>-0.00437</td>
<td>-0.01352</td>
<td>0.00865</td>
<td>0.00495</td>
<td>0.01425</td>
<td>-0.16550***</td>
<td>-0.08868</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02654)</td>
<td>(0.02766)</td>
<td>(0.02493)</td>
<td>(0.02612)</td>
<td>(0.02612)</td>
<td>(0.02612)</td>
<td>(0.02612)</td>
<td>(0.02612)</td>
<td>(0.02612)</td>
<td>(0.02612)</td>
<td></td>
</tr>
<tr>
<td>GNP</td>
<td>33 (PM10)</td>
<td>16 (BS)</td>
<td>-0.0000559'</td>
<td>0.0000149</td>
<td>-0.000395'</td>
<td>0.0000135</td>
<td>0.000075</td>
<td>0.000052</td>
<td>0.000079***</td>
<td>0.000178</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0000317)</td>
<td>(0.0000129)</td>
<td>(0.0000287)</td>
<td>(0.0000171)</td>
<td>(0.0000171)</td>
<td>(0.0000171)</td>
<td>(0.0000171)</td>
<td>(0.0000171)</td>
<td>(0.0000171)</td>
<td>(0.0000171)</td>
<td></td>
</tr>
<tr>
<td>GINI</td>
<td>-1.76</td>
<td>(1.15)</td>
<td>0.03044***</td>
<td>0.02539***</td>
<td>0.02539***</td>
<td>0.02539***</td>
<td>0.03857*</td>
<td>0.01531</td>
<td>-0.15899</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.38)</td>
<td>(1.15)</td>
<td>(1.15)</td>
<td>(1.15)</td>
<td>(1.15)</td>
<td>(1.15)</td>
<td>(1.15)</td>
<td>(1.15)</td>
<td>(1.15)</td>
<td>(1.15)</td>
<td></td>
</tr>
<tr>
<td>DIST</td>
<td>2.08e-06**</td>
<td>(6.91e-07)</td>
<td>1.60e-06**</td>
<td>(6.25e-07)</td>
<td>4.24e-06**</td>
<td>(1.56e-06)</td>
<td>2.10e-13</td>
<td>(2.80)</td>
<td>2.10e-13</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.76)</td>
<td>(3.76)</td>
<td>(3.76)</td>
<td>(3.76)</td>
<td>(3.76)</td>
<td>(3.76)</td>
<td>(3.76)</td>
<td>(3.76)</td>
<td>(3.76)</td>
<td>(3.76)</td>
<td></td>
</tr>
<tr>
<td>DALE</td>
<td>0.01558</td>
<td>-0.01075</td>
<td>0.02333</td>
<td>0.00064</td>
<td>-0.00991</td>
<td>-0.02397</td>
<td>0.0393</td>
<td>0.17084</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02062)</td>
<td>(0.02172)</td>
<td>(0.01817)</td>
<td>(0.01911)</td>
<td>(0.01911)</td>
<td>(0.01911)</td>
<td>(0.01911)</td>
<td>(0.01911)</td>
<td>(0.01911)</td>
<td>(0.01911)</td>
<td></td>
</tr>
<tr>
<td>EDUC</td>
<td>-0.00437</td>
<td>-0.000197</td>
<td>-0.00604</td>
<td>-0.00247</td>
<td>-0.00366</td>
<td>-0.00497</td>
<td>-0.00497</td>
<td>-0.00497</td>
<td>-0.00497</td>
<td>-0.00497</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00598)</td>
<td>(0.00585)</td>
<td>(0.00511)</td>
<td>(0.00499)</td>
<td>(0.00499)</td>
<td>(0.00499)</td>
<td>(0.00499)</td>
<td>(0.00499)</td>
<td>(0.00499)</td>
<td>(0.00499)</td>
<td></td>
</tr>
<tr>
<td>Tau2</td>
<td>0.136***</td>
<td>0.127***</td>
<td>0.110***</td>
<td>0.101***</td>
<td>0.257***</td>
<td>0.278**</td>
<td>0.158**</td>
<td>2.10e-13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4.52)</td>
<td>(4.44)</td>
<td>(4.45)</td>
<td>(4.39)</td>
<td>(2.80)</td>
<td>(2.78)</td>
<td>(2.84)</td>
<td>(2.84)</td>
<td>(2.84)</td>
<td>(2.84)</td>
<td></td>
</tr>
<tr>
<td>Log like.</td>
<td>-67.41</td>
<td>-46.28</td>
<td>-59.21</td>
<td>-37.68</td>
<td>-78.23</td>
<td>-60.82</td>
<td>-56.92</td>
<td>-28.95</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wald chi2</td>
<td>20.84</td>
<td>26.57</td>
<td>21.16</td>
<td>27.52</td>
<td>11.95</td>
<td>9.09</td>
<td>29.86</td>
<td>64.83</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*P<0.1, *P<0.05, ***P<0.01, ****P<0.001, β and se have both been multiplied by 1000.
Table 6.9: Summary of regression results using EB-methodology (sample of lag 0 coefficients (28), lag 1 coefficients (23), and lag 2 coefficients (12)).

<table>
<thead>
<tr>
<th>Equation</th>
<th>Method</th>
<th>Dep. var. OBS</th>
<th>10a β</th>
<th>10b β</th>
<th>11a β</th>
<th>11b β</th>
<th>12a β</th>
<th>12b β</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONSTANT</td>
<td>EB</td>
<td>28 (lag 0)</td>
<td>-3.095*</td>
<td>-2.8248*</td>
<td>-3.2315</td>
<td>-3.0477</td>
<td>6.8799</td>
<td>3.4584</td>
</tr>
<tr>
<td>(se)</td>
<td>EB</td>
<td>28 (lag 0)</td>
<td>(1.5626)</td>
<td>(1.2782)</td>
<td>(3.7424)</td>
<td>(3.7276)</td>
<td>(5.8035)</td>
<td>(4.7874)</td>
</tr>
<tr>
<td>POLL</td>
<td>EB</td>
<td>23 (lag 1)</td>
<td>-0.00426**</td>
<td>-0.00475***</td>
<td>0.00111</td>
<td>0.00055</td>
<td>-0.00871</td>
<td>-0.00795</td>
</tr>
<tr>
<td>(se)</td>
<td>EB</td>
<td>23 (lag 1)</td>
<td>(0.00159)</td>
<td>(0.00136)</td>
<td>(0.00392)</td>
<td>(0.00399)</td>
<td>(0.00573)</td>
<td>(0.00533)</td>
</tr>
<tr>
<td>OVER65</td>
<td>EB</td>
<td>12 (lag 2)</td>
<td>-0.00072</td>
<td>0.01920</td>
<td>0.00888</td>
<td>0.02318</td>
<td>0.00673</td>
<td>-0.04446</td>
</tr>
<tr>
<td>(se)</td>
<td>EB</td>
<td>12 (lag 2)</td>
<td>(0.03982)</td>
<td>(0.03066)</td>
<td>(0.07217)</td>
<td>(0.07606)</td>
<td>(0.18719)</td>
<td>(0.16149)</td>
</tr>
<tr>
<td>GNP</td>
<td>EB</td>
<td>12 (lag 2)</td>
<td>-0.000108*</td>
<td>-1.63e-06</td>
<td>1.64e-07</td>
<td>-8.55e-06</td>
<td>-0.00099</td>
<td>-0.00035</td>
</tr>
<tr>
<td>(se)</td>
<td>EB</td>
<td>12 (lag 2)</td>
<td>(0.000045)</td>
<td>(0.000011)</td>
<td>(0.000053)</td>
<td>(0.000027)</td>
<td>(0.000153)</td>
<td>(0.000056)</td>
</tr>
<tr>
<td>GINI</td>
<td>EB</td>
<td>12 (lag 2)</td>
<td>0.035780***</td>
<td>0.00545</td>
<td>0.01677</td>
<td>0.00760*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(se)</td>
<td>EB</td>
<td>12 (lag 2)</td>
<td>(0.00685)</td>
<td>(0.00685)</td>
<td>(0.01677)</td>
<td>(0.04355)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIST</td>
<td>EB</td>
<td>12 (lag 2)</td>
<td>2.83e-06**</td>
<td>-1.64e-07</td>
<td>5.23e-06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(se)</td>
<td>EB</td>
<td>12 (lag 2)</td>
<td>(9.05e-07)</td>
<td>(1.21e-06)</td>
<td>(3.19e-06)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DALE</td>
<td>EB</td>
<td>12 (lag 2)</td>
<td>0.07652**</td>
<td>0.04667*</td>
<td>0.04342</td>
<td>0.03390</td>
<td>-1.0986</td>
<td>-0.11412</td>
</tr>
<tr>
<td>(se)</td>
<td>EB</td>
<td>12 (lag 2)</td>
<td>(0.02880)</td>
<td>(0.02205)</td>
<td>(0.07168)</td>
<td>(0.07431)</td>
<td>(0.10606)</td>
<td>(0.10456)</td>
</tr>
<tr>
<td>EDUC</td>
<td>EB</td>
<td>12 (lag 2)</td>
<td>-0.01663*</td>
<td>-0.01831*</td>
<td>0.00784</td>
<td>0.00944</td>
<td>0.0220</td>
<td>0.04462*</td>
</tr>
<tr>
<td>(se)</td>
<td>EB</td>
<td>12 (lag 2)</td>
<td>(0.00737)</td>
<td>(0.00565)</td>
<td>(0.01033)</td>
<td>(0.01113)</td>
<td>(0.01724)</td>
<td>(0.02219)</td>
</tr>
<tr>
<td>Tau2</td>
<td>EB</td>
<td>12 (lag 2)</td>
<td>0.1625</td>
<td>9.42e-16</td>
<td>1.5588**</td>
<td>0.1580**</td>
<td>0.191</td>
<td>0.179</td>
</tr>
<tr>
<td>(se)</td>
<td>EB</td>
<td>12 (lag 2)</td>
<td>(1.08)</td>
<td>(0.00)</td>
<td>(2.58)</td>
<td>(2.39)</td>
<td>(1.29)</td>
<td>(1.28)</td>
</tr>
<tr>
<td>Log like.</td>
<td>EB</td>
<td>12 (lag 2)</td>
<td>-61.71</td>
<td>-37.71</td>
<td>-68.44</td>
<td>-49.34</td>
<td>-69.91</td>
<td>-50.62</td>
</tr>
<tr>
<td>Wald chi2</td>
<td>EB</td>
<td>12 (lag 2)</td>
<td>27.00</td>
<td>51.13</td>
<td>8.94</td>
<td>8.94</td>
<td>10.91</td>
<td>11.67</td>
</tr>
</tbody>
</table>

*P<0.1, **P<0.05, ***P<0.01, ****P<0.001, β and se have both been multiplied by 1000.
Table 6. 10: Summary of regression results using EB-methodology (sample consisting of one coefficient (averaged or actual) per country (21), and taking time-period into account (25)).

<table>
<thead>
<tr>
<th>Equation Method</th>
<th>13a</th>
<th>13b</th>
<th>14a</th>
<th>14b</th>
</tr>
</thead>
<tbody>
<tr>
<td>β</td>
<td>β</td>
<td>β</td>
<td>β</td>
<td>β</td>
</tr>
<tr>
<td>OBS</td>
<td>21 (country)</td>
<td>21 (country)</td>
<td>25 (country)</td>
<td>25 (country)</td>
</tr>
<tr>
<td>CONSTANT</td>
<td>-0.3806</td>
<td>-0.3244</td>
<td>-0.5841</td>
<td>-0.5168</td>
</tr>
<tr>
<td>(se)</td>
<td>(1.2077)</td>
<td>(1.2440)</td>
<td>(1.1836)</td>
<td>(1.1926)</td>
</tr>
<tr>
<td>POLL</td>
<td>0.00129</td>
<td>0.00149</td>
<td>0.00065</td>
<td>0.00052</td>
</tr>
<tr>
<td>(se)</td>
<td>(0.00158)</td>
<td>(0.00156)</td>
<td>(0.00144)</td>
<td>(0.00145)</td>
</tr>
<tr>
<td>OVER65</td>
<td>-0.0654*</td>
<td>-0.0598*</td>
<td>-0.0570*</td>
<td>-0.0518*</td>
</tr>
<tr>
<td>(se)</td>
<td>(0.0293)</td>
<td>(0.0335)</td>
<td>(0.02832)</td>
<td>(0.0313)</td>
</tr>
<tr>
<td>GNP</td>
<td>0.000072*</td>
<td>0.000059**</td>
<td>0.000028</td>
<td>0.000042**</td>
</tr>
<tr>
<td>(se)</td>
<td>(0.000033)</td>
<td>(0.000022)</td>
<td>(0.000033)</td>
<td>(0.000016)</td>
</tr>
<tr>
<td>GINI</td>
<td>-0.000072</td>
<td>0.00583</td>
<td>-0.000072</td>
<td>0.00583</td>
</tr>
<tr>
<td>(se)</td>
<td>(0.00860)</td>
<td>(0.00874)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIST</td>
<td>4.00e-07</td>
<td>4.15e-07</td>
<td>4.15e-07</td>
<td>4.15e-07</td>
</tr>
<tr>
<td>(se)</td>
<td>(7.56e-07)</td>
<td>(7.50e-07)</td>
<td>(7.50e-07)</td>
<td>(7.50e-07)</td>
</tr>
<tr>
<td>DALE</td>
<td>0.03015</td>
<td>0.02779</td>
<td>0.03107</td>
<td>0.02508</td>
</tr>
<tr>
<td>(se)</td>
<td>(0.02062)</td>
<td>(0.02426)</td>
<td>(0.01988)</td>
<td>(0.02312)</td>
</tr>
<tr>
<td>EDUC</td>
<td>-0.01579**</td>
<td>-0.01549*</td>
<td>-0.01191*</td>
<td>-0.01076'</td>
</tr>
<tr>
<td>(se)</td>
<td>(0.00601)</td>
<td>(0.00641)</td>
<td>(0.00585)</td>
<td>(0.00618)</td>
</tr>
<tr>
<td>Tau2</td>
<td>0.0666*</td>
<td>0.0698*</td>
<td>0.0667**</td>
<td>0.0662**</td>
</tr>
<tr>
<td>(se)</td>
<td>(2.37)</td>
<td>(2.47)</td>
<td>(2.75)</td>
<td>(2.75)</td>
</tr>
<tr>
<td>Log like</td>
<td>-59.52</td>
<td>-41.00</td>
<td>-56.47</td>
<td>-37.66</td>
</tr>
<tr>
<td>Wald chi2</td>
<td>14.66</td>
<td>14.02</td>
<td>14.33</td>
<td>14.51</td>
</tr>
</tbody>
</table>
6.5.4 Discussion

Given the sensitivity of the findings to the sample choice and the measures used, interpretation is a challenging task. In this section we will first discuss the general findings from regressions carried out on the full sample (sample size 70), the sample excluding the 9 all-cause mortality studies (sample size 61), as well as on the sample consisting of coefficients derived using the TSP measure of particles (sample size 24). We will then compare and contrast these results with those obtained with the samples stratified according to the particle measure and the lag structure, as well as with those that include only one observation (an actual observation or an averaged one) for each country.

Among the moderator variables tested, GINI and the interaction term DIST are the strongest and most stable predictors of the dose-response relationship. The positive coefficients indicate that higher income inequality in general, as measured by the GINI-coefficient, and in high-income countries in particular, as captured by the interaction term DIST, are associated with higher mortality rates from air particles.

The GINI-model yielded a positive coefficient for the GNP term, however, it is only significant in the full sample. In the case of the DIST-model, the GNP term takes a negative coefficient that is weakly significant only in the case of the sample excluding the all-cause mortality studies. This main tendency suggests that the negative direct effect of income is smaller than the indirect increasing effect via its interaction with income inequality. In other words, the effect income has in increasing the positive effect of the inequality in income distribution on the dose-response coefficient is larger than the negative (decreasing) effect of average income on the PM$_{10}$-mortality association.

In section 6.3 we mentioned four possible reasons for income inequality to have an increasing effect on the mortality rate from air pollution. First, decreasing returns to averting and mitigating activities would imply a tendency for higher inequality to be associated with higher mortality rates from changes in air pollution levels. Second, if individuals with relatively low income live in the most polluted areas, and if the adverse effect of air pollution is larger on lower income individuals, then this would once again imply larger PM$_{10}$-mortality in cities with large income inequality. Third, if mortality is associated with stress, and stress is related to social status, then higher income inequality may lead to a larger mortality rate from air pollution. Finally, GINI may have been proxying for inequalities in baseline health or for the quality and availability of health...
care. We did include health services as a moderator variable in one of our regression models, and found a slight decrease in the significance for both GINI and DIST, however, both remained significant. Since our measure of health services was relatively poor, however, this does not necessarily imply that GINI is not proxying for some aspect of health services.

Our findings provide empirical support for the idea that income inequality has an increasing effect on the mortality rate from air pollution. However, with the level of aggregation of our data there is little we can do in terms of choosing among the above explanations.

One implication of the regression results for income level and inequality (as captured in the variable DIST) is rather surprising: inequality in income has a larger positive effect on the PM$_{10}$-mortality association in high-income countries. At the risk of ‘over-interpreting’, two explanations may be given for this phenomenon. First of all, there may be more mitigating and averting activities available in high-income countries to the ones who can afford it. The idea is that the supply side of health investments is more developed in high-income countries, e.g. the existence of low pollution neighbourhoods, well-isolated houses, air filtering systems, medicines, medical advice, high-quality medical procedures. In addition, and related to this, comes the fact that the awareness of the general population, irrespective of income level, about the health risks from air pollution and how to minimise them probably is higher in high-income countries. The second possible explanation for the positive sign of the GNP and GINI interaction term is that high income inequality leads to more psychological stress in a high-income country than in a country with a lower level of average income, or alternatively that the stress induced by inequality is more fatal. First, the higher level of stress may be related to various media incessantly informing and reminding people of the differences. Second, different life-styles (diet, exercise etc.) may mean that stress has a more fatal outcome in high-income countries. Furthermore, high-income countries also tend to have a higher degree of social mobility, something which possibly affects the stress-level experienced by ‘under-achievers’. Finally, there is possibly a stronger focus on material status in high-income countries.

Another slightly surprising finding is the very small and insignificant direct negative effect of income on the dose-response coefficient. It is possible that income does not affect behaviour (averting and mitigating activities) to such an extent that it shows up
very significantly in the regression results. Income level is likely to affect the level of baseline health and education, however these are measured by other moderator variables. A very likely, and perhaps more plausible explanation, however, of the lack of significance of the coefficient for the direct effect of income, is that it is highly correlated with education, disability-adjusted life expectancy, the percentage of the population over 65 years of age, and the ambient level of pollution (see table 6.5). As long as the measures used are not perfect, income may be picking up parts of these effects, some of which work in opposing directions. Alternatively, the other measures may have captured part of the effect that should have been captured by income.

We should also remind the reader that the income variable we ideally would like to have is the average income in the location in question (and preferably of the population particularly vulnerable to air pollution). Since we were not able to obtain this information, we had to use the average income in the relevant country instead. Not only may this imply that our income measure is fairly poor, but in addition it may imply that the GINI coefficient partly proxies for the city income-level.

As for the additional moderator variables POLL, OVER65, EDUC, DALE, and the constant, none was significant in any of the three samples mentioned. The education term was the most consistent of these variables, taking a negative coefficient in all of the samples. The sign suggests that the amount of knowledge or education people have may play a role in avoiding fatal effects from air pollution.

The effect of the average ambient level of pollution (POLL) on the PM$_{10}$-mortality coefficient varies in sign and is insignificant. This finding suggests that non-linearities in the effect of particles on mortality are not very significant. The positive but insignificant coefficient for ambient PM$_{10}$ concentration found in Levy et al. (2000) is in line with our finding.

Similarly, the coefficient for baseline health (DALE) varies in sign from one sample and regression model to another, but it is consistently insignificant. Again, this could indicate that basic health level does not affect significantly the rate at which people die from air pollution. In principle, it is possible that a healthy person could die suddenly from high exposure to air pollution, just as well as an older person or a person suffering from some illness, however, this does not make sense intuitively. The insignificance of the coefficient is most likely due to DALE not being a satisfactory measure of baseline health.
The coefficient for percentage of the population over 65 years of age also varies in sign according to the sample we are analysing, but is consistently insignificant. Only when we excluded any measure of income inequality from our regression equation did OVER65 become significantly negative, something which suggests that it was proxying for income inequality (from table 6.5 we observe that the correlation is -0.40). Levy et al. (2000) found a negative and weakly significant coefficient for population over 65 years of age (per cent) in their regression analysis on a sample of 19 U.S. studies. The difference in results between the two studies may possibly be explained by the fact that this latter study did not include income inequality or baseline health as moderator variables in their regressions.

In section 6.4 we argued that we would expect elderly people to be most at risk from the acute effects of air pollution, and therefore a positive coefficient for this moderator variable. So how do we explain these findings? We suggest that since baseline health is such a multi-faceted concept, no single moderator variable will adequately be able to measure the desired dimension. It is therefore possible that, although we have included DALE as a measure of baseline health in our regression model, OVER65 is still picking up part of the effect of higher baseline health on the PM$_{10}$-mortality association. The reason for this is that individuals tend to survive to a high age if they are in good health, so a society with a large proportion of elderly people may also be a “healthy” society. This will then induce a downward bias in the estimator, which may have cancelled out the expected positive coefficient for OVER65.

The inclusion of a number of moderator variables is required in order to distinguish among potential predictors. However, the lack of independence can pose problems. The findings discussed above of insignificant coefficients, with varying signs, for DALE, OVER65, and POLL may be a function of these correlations. In addition, these moderator variables may not be measuring the desired dimensions or may not represent the actual characteristics of the site. Both DALE and OVER65 are based on country data, rather than data from the site in question. In addition, DALE exists for one year only, and may therefore be very misleading especially in the case of older studies. Although pollution data are available for the particular city we are studying, potentially significant information e.g. on pollutant mixture and particle size was missing.

Interestingly, we find that the estimate of the central effect of PM$_{10}$ on mortality is larger when studies of all-cause mortality are included rather than excluded from the sample.
As the three types of particulate pollution measures are all represented in the 9 studies investigating all-cause mortality, the type of particulate measure used is unlikely to have caused this difference in results. Although the 9 studies may differ from the rest of the sample in some other unidentified way, it is also possible that externally caused mortality is not entirely independent of air pollution. In particular, we suggest that air pollution may affect individuals' level of concentration, stress, as well as visibility and that therefore the amount of externally induced mortality incidences increases with higher levels of air pollution. Furthermore, air pollution will tend to be correlated with the amount of traffic, and thereby possibly with traffic-accidents.

By stratifying the sample according to the particle measure, we observed that the particle measure used affected the findings on the PM$_{10}$-mortality association. Although erroneous conversion factors can explain the differences in central effect estimates according to the pollution measure used, they cannot explain the extent to which the estimated moderator-effects and their significance differ. The fact that most of the moderators are country based means that regressing the observed dose-response coefficients on these variables only makes sense when a variety of countries are included in the sample. Since half of the selected studies using BS to measure air pollution are from Spain and half of those using PM$_{10}$ have been carried out in the US, we argued that the results from these two samples would not be meaningful. For the record, it is worth mentioning that even with samples where such a moderator-effect analysis is meaningful, differences in moderator effects could nevertheless be expected. The studies using BS are likely to differ from those employing PM$_{10}$ in several important respects. First of all, BS is a measure based on older technology and is therefore typically used in earlier air pollution studies. Second, PM$_{10}$ became a commonly used measure of air particles in the US before most other countries, and there is therefore a disproportionate amount of US studies in our sample that use this measure. US studies, as well as newer studies may differ in several respects from older studies and studies carried out in other countries (e.g. in terms of pollutant mixtures, and population characteristics), and this could well be reflected in the moderator coefficients.

Two of the sensitivity tests undertaken in this paper stand out from the rest because of the very different results from those discussed above; the test for sensitivity to the lag structure, and that for the sensitivity to multiple observations for some countries. We will discuss the main findings from these analyses in turn and consider whether they are seriously challenging the results from the rest of the regression analysis.
The results from regressing the dose-response coefficients on the moderator variables varied widely according to the time lag of the pollution measurement used in the original study. For the association between same-day air pollution and mortality we found that the GNP coefficient is significantly negative and the DIST-coefficient significantly positive. When GINI is included instead of DIST, GNP retains the negative sign but is insignificant. These results are in line with our previous findings. However, for air pollution lagged one day there is no significant effect from income, income inequality, or their interaction term. In the case of a two-day lag, the effects are similar to the same-day effects for the income-related variables but the significance levels are lower, and only GINI is weakly significant.

The widely varying results according to the lag structure are most likely due to the small sample sizes. The sample for lag zero consists of 28 observations, for lag one 23, and in the case of lag two, 12. As discussed in sub-section 6.5.1, one has to be extremely cautious with results based on small sample sizes, and in particular in the case of the smaller samples, i.e. lag one and two, the size of the samples is likely to drive the results. The smaller the size, the wider the confidence interval, reflecting an amount of uncertainty about the parameter being estimated, which becomes undesirable at some point. Everything else equal, the results from the analysis upon the largest sample are more likely to have some meaning, although caution is still required. These results were found to be in line with our previous findings.

If we have included the main explanatory variables in our regression model, then the representativeness of the sample would not have been an issue. The reason for this is that the requirement for a sample to be representative is based on the idea that there may be underlying factors not known or fully understood by the researcher that influence the relationship under investigation. If, however, the moderators cannot adequately account for the discrepancies between study findings, then the fact that the sample may be unrepresentative constitutes a challenge. How serious this challenge is depends to some extent on what we would like to do with our findings. In particular, the challenge will be present if we would like to generalise our results in order to tailor-make dose-response coefficients to different circumstances. If, on the other hand, the purpose of the meta-analysis is first and foremost the wish to investigate whether various variables are significantly affecting the PM$_{10}$-mortality ratio, rather than the actual size of the effect, the unrepresentativeness of the sample may not be such a problem.
It is, however, not straightforward to determine what would have been a representative sample—'the concept of representativeness has always been a somewhat ambiguous idea, but the concept is useful in helping to illuminate potential problems in drawing inferences from samples' (Cooper and Hedges (1994), p. 35). Like Levy et al. (2000), we assume that the studies are independent samples from a random distribution of the conceivable population of studies. The latter being all the potential studies that could have been carried out given our inclusion and exclusion criteria. This assumption is clearly difficult to support, mainly due to the overweighing of certain countries and cities, but also due to the sample entailing multiple studies conducted by the same author(s) and using the same methodology. On a more general note, there are two main reasons why exhaustiveness of sampling does not necessarily yield a representative sample of the universe. First, some types of studies in the intended universe may not have been conducted (e.g. because of unreliable data), and second, although studies may have been conducted, they may not have been reported in journals or other forums accessible to the researcher.

To determine whether the results were sensitive to study selection, and in particular to the inclusion of multiple observations from certain countries, we therefore conducted analyses where each country entered with only one observation, either an averaged one or from one single study (randomly chosen). Averaging the data over several studies, and using the average as if it were an observation is a very dubious method, and is likely to create correlations between variables where these do not exist. These problems are worsened in the case where the averaged data came from studies covering very different time periods. Including single studies is therefore a preferred method when testing the sensitivity to multiple country observations. The main objections to this latter method are, first, that the sample becomes smaller, second, that we are ignoring/rejecting information that we actually have, third, that the observation that happen to be included may be unrepresentative for the country in question, and finally, the resulting sample may not necessarily be more representative than the full sample. Compared to the results from the full study, this country-weighted sensitivity analysis found a more significant effect of education in decreasing the particle-mortality effect. Furthermore, no significant role was found for either the GINI-coefficient or its interaction term with GNP. The role of income appears relatively unchanged, with a positive and significant effect in the GINI-model, and no significant effect in the DIST-
model. These findings are similar to those obtained on the sample where data for the same country, of the same time-period, were averaged.

When comparing the result using the full sample with those using the ‘one-country-one-observation’ sample, the differences in results are of degree rather than of tendency/direction. The coefficient for education is consistently negative, for income inequality it is consistently positive, and as for income, it tends to have a direct effect which is either negative or zero, and an indirect effect via income inequality which is positive. Significance varies between the samples.

We have seen that there are weaknesses related to both samples. We nevertheless propose that the results based on the full sample are the preferred ones, since the sample with one observation for each country excluded existing information without necessarily becoming more representative of the universe of potential studies (given our selection criteria).

Since the inter-study variability remained significant, it is quite likely that we omitted relevant moderators that could have explained at least part of it. It is also possible that omitted relevant moderators can explain the significance of some of the associations we found, e.g. of income inequality. Possible omitted relevant explanatory variables are smoking habits, the ratio of fine to coarse particles, other pollutants, central air conditioning prevalence, and cooking and heating mechanisms used. Furthermore, the data used to capture the moderator variables may have been poorly suited for that purpose. Most of the data were given on a country-wide level, whereas the studies were carried out for a specific region or city within that country which may well differ significantly from the country-average with respect to that particular characteristics. In addition, some of the moderator concepts were too complicated and vague to be captured by one type of indicators, e.g. health services.

Even if the above weaknesses could be solved, the problem of selection bias remains. Although we included all the available studies that fulfilled the selection criteria set out in section 6.4, there is no way of avoiding possible “publication-bias” or “submission-bias” (i.e. studies that do not find a significant relationship between PM$_{10}$ and mortality, or that find a negative association, are not published or submitted for publication). Nevertheless, the file-drawer bias is highly unlikely to change the finding of a significant PM$_{10}$-mortality link. With the assembled sample (70) of largely positive estimates, 70 unpublished estimates with central estimates of −0.5 per cent and the mean
variance of our 70 study estimates would be required to yield a statistically insignificant
grand mean estimate. A further source of selection bias lies in the representativeness of
our study locations. As argued earlier, it is possible that the cities in our analysis are not
representative of the true range of conceivable studies. As argued by Levy et al. (2000), it
is possible that a subset of cities “that were more or less prone to PM health effects were
chosen for epidemiologic analyses” (p.115). The main problem with this challenge to our
sample is to know the range of conceivable studies. Nevertheless, the cities and countries
included in our sample represent a wide range of population characteristics and pollution
concentration levels.

One particular and potentially important weakness of the original studies may
have influenced the various findings from our meta-analysis. As discussed in chapter 5 of
this thesis, and briefly mentioned towards the end of section 6.4.1, measuring PM_{10}
exposure with readings of ambient pollution from central monitoring stations may be
highly problematic, especially in the case of developing countries. If indoor air pollution
in developing countries, caused by the use of highly polluting heating and cooking fuels
and little ventilation, is a larger health risk than outdoor air pollution, then the exposure-
response association may be much larger but not be revealed by studies using exposure
measurements obtained from central monitoring stations. The individuals vulnerable to
outdoor air pollution may already have died from indoor air pollution, and the coefficient
will not capture these deaths adequately as long as indoor and outdoor air pollution do
not strongly covary. A possible implication of this measurement error is that those
moderator variables that tend to take a significantly different value in the case of
developing than developed countries, i.e. income per capita, education, and disability-
adjusted life expectancy may have lost significance.

A number of more general weaknesses and limitations to our meta-analysis are
necessary to point out. First of all, it is important to note that as the meta-analysis uses a
correlational design for its moderator variables, no causal inference can appropriately be
made. Secondly, when the moderator variables cannot adequately account for the
discrepancies between study findings, as is the case for our full sample, it becomes
unclear whether the various research operations represent a common underlying
construct, and uncertainty in the empirical effects of the moderators emerges.
Furthermore, unresolved inconsistencies between studies may compromise the ability to
detect true relations among variables. ‘In such cases’ states Cooper and Hedges (1994)
meta-analytic findings should be accorded relatively low certainty’ (p.490). Thirdly, even if the moderator variables successfully account for the empirical variability in study outcomes, the proper interpretation of the relation may remain an issue, especially when the moderators are identified on a between-study basis. The moderator is then an attribute of an entire study, and individuals in the primary research have not been randomly assigned to levels of the moderator variable. Such moderator variables may be confounded with other variables that covary with it.

Our first study-objective was to increase our understanding of what affects the amount of deaths that are related to air pollution. From our regression results we have seen that several of the population characteristics included as moderator variables may play a significant role in the association between particle pollution and mortality. In particular, income inequality may be associated with a larger effect of a change in pollution on the mortality rate, especially in high-income countries. The direct effect of income may be slightly decreasing the effect of pollution on mortality, and a similar finding was made for education.

Second study-objective was to help highlight areas where further studies may be needed. Indeed, we have pointed out that the discrepancies between study effects remain significant, that some of the data are not very well-suited to capture the moderator effects in question, that no causal inferences can appropriately be made for the moderator variables, that the sample of studies may not be representative of the hypothetical universe of studies, and that there may be serious measurement errors in the original studies. Subsequent research will need to address all of these concerns. The first point can be addressed by adding other possible moderator variables that have been discussed in the literature (e.g. the ratio of fine to coarse particles, other pollutants, gas stove prevalence, central air conditioning prevalence, warm air heating prevalence, smoking habits) to the ones discussed in the present paper. This may not only eliminate the discrepancies between study-effects, but may also alter the findings on the already included moderators. The second point can be approached by for example trying to obtain city- or location-specific information on the relevant moderator variables, and by more generally encouraging the primary studies to gather and publish this information. Meta-analysts should also be on the lookout for improved indicators of the moderator effects. In order to be able to make causal inferences, the third point, two criteria in addition to covariance have to be fulfilled; temporal ordering (i.e. that the cause should temporally precede the effect) and
isolation (that the effect was not caused by something else). This suggests that future PM$_{10}$-mortality research could be directed towards panel-studies gathering individualised time-series data on the relevant moderator variables. As for the latter two points, future primary research would benefit from being directed towards the understudied populations; in particular, more studies from poor countries are required, and would also benefit from improved exposure measurements; in particular, including indoor air pollution exposure.

Final study-objective was, through the ensuing coefficients of the moderator variables, to help transferring the dose-response coefficients to countries where empirical studies have not yet been feasible or to forecast the effects of policies targeting air pollution. The coefficients derived from our meta-analysis could in theory help us tailor-make dose-response coefficients to local conditions. Indeed, for an example of how this can be done refer to Maddison and Gaarder (2001). However, as long as the between-study variation remains significant, we suggest proceeding with caution. Rather than using the exact coefficients derived in our meta-analysis, and which may change significantly once the whole variation has been explained by additional moderators, we propose using the findings to suggest relative differences. If we for example have estimated the dose-response coefficient for one city, and would like to transfer it to another city that only differs in that it has a higher income inequality, we suggest that the empirical coefficient at hand probably will constitute a lower bound of the actual effect. If only the education level between two cities differed, we could again propose that the dose-response coefficient would probably be higher in the city with lower levels of education. A further note of caution is however in order. For any city or location with a value for one of the moderator variables outside the range covered in the meta-analysis, transfer will strictly speaking not be valid.
6.6 Conclusion

In this paper we have applied the Empirical Bayes mixed-effect model to the largest sample of air-pollution mortality studies to date, from the widest range of countries, in order to determine whether the variability in effect estimates can be explained by a selected group of population characteristics. Theoretical plausibility, novelty, and availability of data guided the selection of study-characteristics included as potential moderator variables.

We estimated that a 100-µg/m³ increase in PM₁₀ concentrations was associated with a 6 per cent increase in daily mortality, which was in line with previous findings, although slightly lower. When including the moderator variables in our analysis we found that higher income inequality in general, as measured by the GINI-coefficient, and in high-income countries in particular, is associated with higher mortality rates from air particles. For our full sample, a four-point increase in the GINI-coefficient leads to an increase in the mortality rate from a 100-µg/m³ increase in PM₁₀ of almost 1 per cent (i.e. from 6 to 7 per cent). We further found some evidence that the PM₁₀-mortality relationship may be stronger in locations with lower education. Furthermore, there appears to be a positive (increasing) effect of income on the association between income inequality and the dose-response relationship and a negative direct effect of GNP per capita on the dose-response coefficient, with the former outweighing the latter. Our findings were quite sensitive to the samples and measures analysed in our sensitivity analyses.

Although our interpretations of these associations are highly tentative, we nevertheless find it useful to summarise potential explanations for the above findings. In addition to income inequality possibly proxying for the quality and availability of health services, or health inequalities, larger income inequalities may be associated with stress, and stress may in turn affect mortality from air pollution. Furthermore, decreasing returns to averting and mitigating activities would imply a tendency for higher inequality to be associated with higher mortality rates. The fact that income inequality increases the effect of pollution on mortality more in high-income countries may, firstly, be due to higher availability of mitigating and averting activities in high-income countries to the ones who can afford it, and secondly to inequality creating relatively more psychological stress in high-income countries. The latter could plausibly be due to the role of media, higher social mobility, and stronger focus on material status in higher income countries. Alternatively, the stress induced by inequality may be more fatal in high-income
countries due to different life-styles. Finally, it is not unlikely that the income level of the relevant population does not correspond to the per capita income level of the whole country – the combination of income and income inequality may then together proxy for this variable.

We argue that the level of education affects the knowledge people have about health, health production, and the connection between air pollution and health, and hence the amount of averting and mitigating expenditures undertaken and the efficiency of these expenditures. As for the generally small (and mostly insignificant) negative direct effect of income on the dose-response association, we mention the possibility that income does not affect behaviour to such an extent that it shows up very significantly in the regression results. Income level is likely to affect the level of baseline health and education, however these are measured by other moderator variables.

Although baseline health, the percentage of people over 65 years of age, and the level of ambient particle concentration were not found to play significant roles, this may be due to the quality and level of aggregation of available data, as well as the correlation between several of the moderator variables. Note also that even though we have invoked a number of moderator variables, and some of them are significant, substantial unexplained variation in the coefficients remains. This suggests that aspects of the relationship between air pollution and mortality remain unexplained or that they are due to any of a number of variables that we have mentioned but for which we did not have adequate data. Our findings on the included moderator variables must be qualified by these considerations.

We have also argued that several of the moderator variables (in particular income, education, and baseline health) may have lost significance due to measurement error in the original studies. Readings of ambient pollution from central monitoring stations may be particularly unsuited as measurements for PM$_{10}$ exposure in developing countries. The individuals potentially vulnerable to outdoor air pollution may already have died from indoor air pollution, and the coefficient will not capture these deaths as long as indoor and outdoor air pollution do not covary.

Interestingly, we find a slight tendency for the effect of pollution to be larger in studies looking at all-cause mortality than those excluding deaths from ‘external’ causes, potentially implying that externally caused mortality is not entirely independent of air pollution. Furthermore, we find the estimated moderator effects to differ depending on the
lag-structure, opening up for the possibility that behaviour and population characteristics affect the timing of mortality.

The paper has introduced a variety of population characteristics not previously investigated, and uncovered that a number of these have a measurable influence on the magnitude of the PM-mortality relationship. The ensuing coefficients can in the first instance be used to suggest whether a dose-response coefficient transferred from one country to another is likely to under- or overestimate the mortality caused by ambient particles. When additional moderator variables are included that successfully explain the between-study variability, the ensuing coefficients can be used to tailor-make dose-response coefficients to local conditions. Subsequent EB analysis may benefit from adding other possible moderator variables to the ones discussed in the present paper. In addition, it should focus on obtaining more location-specific information on the relevant moderator variables and improved indicators of the moderator effects. Future meta-analyses would benefit from primary research being directed towards understudied areas and populations, such as in developing countries and rural areas, and from improving exposure measurements (e.g. increase the number of monitoring stations, include indoor air pollution exposure). We also recommend future PM$_{10}$-mortality studies to explore the possible causal inferences to be made. For these purposes, the primary research will need to be directed towards panel-studies gathering individualised time-series data on the relevant moderator variables.
REFERENCES


Aunan, K., G. Patzay, H. A. Aaheim, and H. M. Seip (1997), Health and environmental
benefits from the implementation of an energy saving program in Hungary, CICERO Report, No. 97/1, University of Oslo, Oslo.


No. 7, pp. 494-501.


Dockery, D.W., C.A. III Pope, X. Xu, J.D. Spengler, J.H. Ware, M.E. Fay, B.G. Ferris, 258


Environmental Protection Agency (1996), *Air quality criteria for particulate matter*, Volumes I – III, EPA/600/P-95/001cF, United States.


Freeman, A.M. (1993), *The measurement of environmental and resource values. Theory*


Gwatkin, R. (1999), *Poverty and inequality in health within developing countries*, paper prepared for the Ninth Annual Public Health Forum of the London School of Hygiene and Tropical Medicine.


Halvorsen, B. (1997), *Methodological issues in applying stated preferences data to the valuation of public goods*, Økonomiske Doktoravhandlinger, No. 39, University of
Oslo.


National Research Council Board on Environmental Studies and Toxicology Committee on Advances in Assessing Human Exposure to Airborne Pollutants (1991), Human exposure assessment for airborne pollutants advances and opportunities, National Academy of Sciences, Washington D.C.


Schwartz, J. and A. Zanobetti (2000), 'Using meta-smoothing to estimate the dose-response trends across multiple studies, with application to air pollution and daily deaths', *Epidemiology*, Vol. 11, No. 6, pp. 666-672.


272


Appendices

Appendix 3.A: Mathematical Notes

A1: Envelope Theorem

We know that the optimal choice function $L(w)$ must satisfy the condition

$$
\frac{\partial^2 f(L(w), w)}{\partial L \partial w} \equiv 0.
$$

Differentiating both sides of this identity gives:

$$
\frac{\partial^2 f(L(w), w)}{\partial^2 L} \frac{dL(w)}{dw} + \frac{\partial^2 f(L(w), w)}{\partial L \partial w} \equiv 0.
$$

Solving for $\frac{dL(w)}{dw}$, we have

$$
\frac{dL(w)}{dw} = -\frac{\frac{\partial^2 f(L(w), w)}{\partial L \partial w}}{\frac{\partial^2 f(L(w), w)}{\partial^2 L}}.
$$

The denominator of this expression is negative due to the second-order condition for maximisation. Noting the minus sign preceding the fraction, we can conclude that the sign of $\frac{dL(w)}{dw}$ equals the sign of $\frac{\partial^2 f(L(w), w)}{\partial L \partial w}$.

A2: Releasing Assumption of Additive Separability

If we do not make the assumption of additive separability we obtain the following differential of the first-order condition in equation (3.12):

$$
\frac{\partial}{\partial w} \left[ U_c(w\Phi(L)-a, L)w\Phi'(L) + U_L(w\Phi(L)-a, L) \right] =
$$

$$
U_c(w\Phi(L)-a, L)\Phi'(L) + U_{cc}(w\Phi(L)-a, L)w\Phi'(L) + U_{lc}(w\Phi(L)-a, L)\Phi(L) =
$$

$$
U_c(w\Phi(L)-a, L)\Phi'(L) \left[ 1 - \frac{\Phi(L)w}{\Phi(L)w-a} \right] - \frac{cU_c}{U_e} \left( \frac{U_L}{U_e} + \frac{\Phi(L)}{\Phi'(L)} \right).
$$

This expression will have the same sign as (see A1):

$$
- \left[ 1 - \frac{\mu}{\sigma} \left( \frac{U_L\Phi(L)}{U_L\Phi'(L)} \right) \right]
$$

The marginal effect of efficient leisure on optimal amount of efficient labour, $\Phi'(L)$, is negative, whereas it seems reasonable to assume that the marginal utility of efficient
leisure is increasing in consumption, and hence that $U_{LC}$ is positive. In this case, using a non-additively separable utility function implies that $\frac{dL(w)}{dw}$ will have a positive sign for a higher elasticity of substitution or a lower net minimum consumption requirement ($a$) than in the example discussed in the text.

A3: Derivation of Expression (3.13)

$$\frac{\partial}{\partial w} \left[ U_c \left( w\Phi(L) - a, L \right) \right] = U_c \left( w\Phi(L) - a, L \right) + U_{\infty} \left( w\Phi(L) - a, L \right) \Phi(L) \Phi'(L)$$

$$= U_c \left( w\Phi(L) - a, L \right) \Phi'(L) \left[ 1+ \frac{U_{\infty} \Phi(L) w}{U_c} \right] = U_c \left( w\Phi(L) - a, L \right) \Phi'(L) \left[ 1- \left( \frac{\Phi(L)w}{\Phi(L)w-a} \right) \left( \frac{cU_{\infty}}{U_c} \right) \right]$$

(Note: $c = \Phi(L)w - a$). Since $U_c$ is positive and $\Phi'(L)$ is negative, the expression above will be of the same sign as expression (3.13) in the text, i.e.:

$$- \left[ 1- \left( \frac{\Phi(L)w}{\Phi(L)w-a} \right) \left( \frac{cU_{\infty}}{U_c} \right) \right]$$

A4: Alternative Utility Functions

First, assume consumption and leisure to be perfect substitutes, e.g.:

$$U(c, L) = \alpha (w\Phi(L) - a) + \beta L$$

Now what we want is the sign of $dL(w)/dw$, which will have the same sign as the expression:

$$\frac{\partial}{\partial w} \left[ \alpha w\Phi'(L) + \beta \right] = \alpha \Phi'(L) < 0$$

We know from before that $\Phi'(L) < 0$ and hence the sign of $dL(w)/dw$ is negative.

Next we let consumption and leisure be perfect complements, e.g.:

$$U(c, L) = \min (\alpha (w\Phi(L) - a), \beta L).$$

In this case the derivative of the first order condition with respect to wage is the following:

a) For $\alpha (w\Phi(L) - a)$ being the minimum of $\min (\alpha (w\Phi(L) - a), \beta L)$: $\alpha \Phi'(L)$

b) For $\beta L$ being the minimum of $\min (\alpha (w\Phi(L) - a), \beta L)$: $=0$

The sign of $dL(w)/dw$ in case a) is negative, whereas in case b) wage has no effect on leisure.

Let us now look at quasilinear preferences between consumption and leisure, e.g.:
\[ U(c, L) = \nu(w \Phi(L) - a) + L \]

In this case the derivative of the first order condition with respect to wage is:

\[
\frac{\partial}{\partial w} \left[ \frac{1}{2} (w \Phi(L) - a)^{\frac{1}{2}} w \Phi'(L) + L \right] = \\
- \frac{1}{4} (w \Phi(L) - a)^{\frac{3}{2}} \Phi(L) w \Phi'(L) + \frac{1}{2} (w \Phi(L) - a)^{\frac{1}{2}} \Phi'(L) = \\
\frac{1}{2} (w \Phi(L) - a)^{\frac{1}{2}} \Phi'(L) \left[ 1 - \frac{\Phi(L) w}{2(w \Phi(L) - a)} \right]
\]

The first part of the expression above (before the term in brackets) is negative as long as consumption is positive. Then \( \frac{dL(w)}{dw} \) will be of a positive sign if \( a \frac{a}{w} > \frac{\Phi(L)}{2} \), and of a negative sign otherwise.

For the opposite quasi-linear indifference curve, e.g. \( U(c, L) = \nu(w \Phi(L) - a) \), we have:

\[
\frac{\partial}{\partial w} \left[ \frac{1}{2} L^{\frac{3}{2}} + w \Psi(L) \right] = \Psi(L)
\]

Hence \( \frac{dL(w)}{dw} \) is negative.

As we see the results are very sensitive to the type of utility function we use. It seems, however, unlikely that consumption and leisure should be perfect substitutes or perfect compliments. In that case, \( \frac{dL(w)}{dw} \) can take both a positive and a negative sign.

Appendix 3.B: Potentially Different Productivity in Work and Leisure

The rather restrictive assumption made in this paper that the marginal productivity of time in work and leisure is the same is here relaxed. Assume that the illness reduces the value of time, in terms of productivity in work and leisure, but not necessarily equally much. Each hour of \( r \) that is spent recuperating is regarded as equivalent to \( s \), \( 0 < s < 1 \) hours of leisure. Whereas time in the recuperation phase that is spent working (i.e. \( r(1-s) \)) is done so at productivity level \( p \), \( 0 < p < 1 \). The rest of the time, \( l - r \) can also be spent in work or leisure. Let each hour of leisure be equivalent to \( \sigma(x) \), \( 0 \leq \sigma(x) \leq l \) hours of effective leisure and each hour of work have productivity \( \pi(x) \), \( 0 \leq \pi(x) l \).

Assume that if the entire recuperation phase is spent recuperating, then the post-recuperation phase gives full productivity in work or leisure, i.e. \( \pi(l) = \sigma(l) = l \).
whereas if the recuperation time is spent working the productivity in the post-
recuperation time will be the same as during recuperation, i.e. \( \pi(0) = p, \sigma(0) = s \).

For any given choice of \( x \) and \( z \) the consumer will get efficient leisure

\[
L = s \, r \, x + \sigma(x) \, (1-r) \, z
\]

This takes the minimum value \( L = 0 \) when \( x = z = 0 \), and a maximum value \( L = 1-r + \sigma \cdot s \).

Associated with this choice of \( x \) and \( z \) is the amount of work:

\[
H = p \, r \, (1-x) + \pi(x)(1-r)(1-z)
\]

The maximisation problem is now defined in the following way:

\[
\Phi(\xi) \equiv \max_{x \geq 0, z \geq 0} pr(1 - x) + \pi(x)(1 - r)(1 - z) \quad \text{s. t. } rsx + \sigma(x)(1-r)z \geq L, x \leq 1, z \leq 1
\]

\[
(3.1')
\]

This gives the following first order conditions (FOCs):

\[
\frac{\partial \Phi}{\partial x} : (\lambda \, s - p) \, r + (1 - r) \, (\pi'(x) \, (1 - z) + \lambda \, \sigma'(x) \, z) \leq \mu, \quad x \geq 0 \quad [\partial \Phi/\partial x]x=0
\]

\[
x \leq 1, \quad \mu \geq 0
\]

\[
\frac{\partial \Phi}{\partial z} : (\lambda \, \sigma(x) - \pi(x)) \, (1 - r) \leq \nu
\]

\[
z \leq 1, \quad \nu \geq 0
\]

\[
\frac{\partial \Phi}{\partial \lambda} : s \, r \, x + \sigma(x) \, (1 - r) \, z \geq L
\]

\[
\lambda \geq 0
\]

\[
(3.2')
\]

If we first analyse the case where \( z > 0 \), we can distinguish between two types of
regimes:

**Regime 1a:** This is the case where the entire post-recuperation phase is spent in leisure,
i.e.: \( z = 1 \) and hence \( \nu \geq 0 \) and \( [\partial \Phi/\partial z]=0 \). Plugging these values into the second FOC,
this implies that \( \lambda > \pi(x) / \sigma(x) > 0 \), and from the first FOC this value for \( \lambda \) implies
that \( \mu \geq 0 \) and \( x \leq 1 \). By substituting these values for recuperation and post-recuperation
leisure, \( x \) and \( z \), into equation (3.1') and (3.2') we obtain the following expressions for
efficient leisure and labour:

\[
L = s \, r \, x + \sigma(x)(1-r)
\]

\[
\Phi(L) = r(p - px + sx) + (1-r) \, \sigma(x) - L; \quad \Phi'(L) = -1
\]

**Regime 1b:** Next, we look at the case where some of the post-recuperation phase is spent
in leisure, i.e.: \( 1 > z > 0 \). This implies that \( \nu = 0 \) and \( [\partial \Phi/\partial z]=0 \), and hence \( \lambda = \pi(x)/\sigma(x) > 0 \) (from second FOC). From the first FOC we find that this constraint on \( \lambda \) in turn
implies \( \mu \geq 0 \) and \( x \leq 1 \). Efficient leisure and labour are then given by:

\[
L = s \, r \, x + \sigma(x)(1-r) \quad \{< s \, r \, x + \sigma(x)(1-r) \}
\]

\[
\Phi(L) = r(p - px + sx) - (1-r)(\pi(x)(1-z) + \sigma(x)z) - L;
\]

\[
\Phi'(L) = -1
\]

280
In these regime 1 solutions we see that only when $\lambda \geq \frac{p}{s}$ will recuperation be spent entirely in leisure (i.e. $\mu > 0; x = 1$). In this case efficient labour will in both cases be given by:

$$\Phi(L) = sr + (1 - r) - L$$

**Regime 2:** The second type of regime occurs when: $0 < x < 1$, which implies $\mu = 0$ and $[\partial \Phi / \partial x] = 0$. Substituting these values into the first FOC we then find $\lambda < \frac{p}{s}$. In this case $z = 0$ only if $\lambda < \frac{\pi(x)}{\sigma(x)}$. Assuming this is the case, the solution may be usefully split into two scenarios:

**Scenario 1:** In the case where $\frac{\pi(x)}{\sigma(x)} > \lambda > 0$, i.e. where there is still some efficient labour to be gained by trading in efficient leisure time, we obtain the following expression for leisure from the third FOC: $L = sr x < s r$. Efficient labour, and the first and second derivative with respect to leisure can be expressed as:

$$0(L) = r(p - px + sx) + \frac{1 - r}{sr} \frac{L}{sr} - \lambda < 0; \quad 0''(L) = \frac{1 - r}{(sr)^2} \frac{\pi''(L)}{sr} < 0.$$  

This implies that the budget constraint is concave and that $\pi'\left(\frac{L}{sr}\right) < \frac{sp}{1 - r}$.

**Scenario 2:** In the case where $\frac{\pi(x)}{\sigma(x)} > \lambda = 0$ (i.e. no additional efficient labour can be gained by trading in efficient leisure) leisure is given by the following inequality: $L < s r x$. The slope of the efficient time locus is $\Phi'(L) = -\lambda = 0$. Hence, for a certain amount of recuperation time, $x$, the efficient labour that can be gained by working an additional hour during the recuperation time is equal to that which can be gained by recuperating that hour and thus gaining in post-recuperation productivity, i.e.: $\pi'(x) = \frac{sp}{1 - r}$. Knowing that $x > \frac{L}{sr}$, and that there are decreasing returns to recuperation we can further deduce the following: $\pi'\left(\frac{L}{sr}\right) > \frac{sp}{1 - r}$.

In this more general discussion where productivity in work and leisure are not necessarily the same, we find that individuals will work during part of their recuperation time only when the shadow price of leisure is smaller than the ratio of work to leisure productivity during recuperation (i.e. $\lambda < \frac{p}{s}$). However, working during recuperation does not necessarily imply that no leisure is taken in the post-recuperation phase. In
particular, this latter case will only arise if the shadow price of leisure is smaller than the ratio of work to leisure productivity during the post-recuperation period (i.e. $\lambda < \pi(x) / \sigma(x)$). On the other hand, if the shadow price is equal to or larger than both these ratios then recuperation will be spent entirely in leisure and leisure will also be taken in the post-recuperation phase.

As long as the ratio of the productivity in work to the productivity in leisure is the same during the recuperation period as during the post-recuperation period, something which seems a likely and reasonable assumption, the relaxation of the assumption of equal productivities does not change the model insights in any significant way.

**Appendix 3.C: Critical Wage Derived in the Cases Where $\sigma = 1$ and $\sigma = 0.5$**

Analytical expressions for the critical wage can be derived in at least two cases, the first being the Cobb-Douglas case where $\sigma = 1$, and the second is where $\sigma = 0.5$.

Assuming that $L \leq T / (1 + \gamma)$, the critical wage when $\sigma = 1$ can be expressed as follows:

$$\bar{w} = \frac{a}{T - (1 + \gamma)L} = \frac{a}{1 - r - sr \frac{\alpha}{1 - \alpha}}$$

$$\frac{\partial \bar{w}}{\partial \alpha} = \frac{1}{T - (1 + \gamma)L} > 0, \quad \frac{\partial \bar{w}}{\partial \alpha} = \frac{aL}{(1 - r - L \frac{\alpha}{1 - \alpha})^2} > 0$$

with

$$\frac{\partial \bar{w}}{\partial s} = \frac{ar \frac{\alpha}{1 - \alpha}}{(1 - r - L \frac{\alpha}{1 - \alpha})^2} > 0, \quad \frac{\partial \bar{w}}{\partial \gamma} = \frac{a(1 + s \frac{\alpha}{1 - \alpha})}{(1 - r - L \frac{\alpha}{1 - \alpha})^2} > 0$$

The critical wage $\bar{w}$ is strictly increasing in $\alpha$, and thus strictly decreasing in unearned income, $\bar{y}$. The critical wage is further increases with sickness duration, $r$, productivity during sickness, $s$, as well as with the weights given to consumption in the preference function, $\alpha$.

Turning to the case where $\sigma = 0.5$, and substituting this elasticity into equation (3.19) a quadratic function in $\sqrt{w}$ is obtained, and the solution is the following:

$$\bar{w} = \left[ \frac{T \gamma \sqrt{(L \gamma)^2 + 4a(T - L)}}{2(T - L)} \right]^2$$

For completeness, let us briefly look at the effect of unearned income on leisure choice and on the effect of wages upon leisure.

Regime 1: Differentiating the expression for leisure (3.19), and the expression for the effect of wage upon leisure (3.20) with respect to unearned income we obtain:

\[
\frac{dL}{d\gamma} = \frac{1}{\gamma w^\sigma + w} > 0, \quad \frac{dL}{dw} = -\frac{\left(1 + \gamma \sigma w^{\sigma-(1-\sigma)}\right)}{\left[\gamma w^\sigma + w\right]^2} < 0
\]

Leisure will therefore be increasing in unearned income, ceteris paribus, whereas the positive effect of wage upon leisure is decreasing with higher unearned income.

Regime 2: It can furthermore be seen in equations (3.21), (3.22), and (3.23) that both the existence wage and the critical wage are decreasing in unearned income.
Appendix 4.A: Differentiations

A.1

Differentiations of equation (4.8):

\[ AIL_1 = \psi wr(1 - s) \]

\[ \frac{\partial AIL}{\partial w} = \frac{\psi w r(1 - s) \left( \frac{\psi w^\sigma + \sigma w}{\psi w^\sigma + w} \right)}{\left( \frac{\psi w^\sigma + w}{\psi w^\sigma + w} \right)} = \frac{r(1 - s)\psi (1 + e_x)}{\psi (1 - \sigma)(1 - \psi)} > 0, \]

\[ \frac{\partial AIL}{\partial r} = \psi w (1 - s) > 0, \quad \frac{\partial AIL}{\partial s} = -\psi w r < 0, \quad \frac{\partial AIL}{\partial a} = 0, \quad \frac{\partial AIL}{\partial \alpha} = \frac{\psi(1 - \psi)\sigma w r (1 - s)}{\alpha(1 - \alpha)} > 0, \]

\[ \frac{\partial AIL}{\partial \sigma} = \frac{\psi w^\sigma r (1 - s) \left[ \log(\alpha w) - \log(1 - \alpha) \right]}{\psi w^\sigma + w} \left[ 1 - \frac{\psi w^\sigma}{\psi w^\sigma + w} \right] = \psi (1 - \psi) w r (1 - s) \left[ \log(\alpha w) - \log(1 - \alpha) \right] = 0, \]

Differentiations of equation (4.8'):

\[ AIL_1 = \alpha wr(1 - s) \]

\[ \frac{\partial AIL}{\partial w} = \alpha (1 - s) > 0, \quad \frac{\partial AIL}{\partial r} = \alpha (1 - s) > 0, \quad \frac{\partial AIL}{\partial s} = -\alpha w r < 0, \quad \frac{\partial AIL}{\partial a} = 0, \quad \frac{\partial AIL}{\partial \alpha} = w r (1 - s) > 0, \]

Differentiations of equation (4.9):

\[ RIL = \frac{\psi wr(1 - s)}{\psi (w - a) + a} \]

\[ \frac{\partial RIL}{\partial w} = \frac{r(1 - s)\psi \left[ \sigma (1 - \psi) \right]}{(\psi (w - a) + a)^2} > 0, \]

\[ \frac{\partial RIL}{\partial r} = \frac{\psi w r(1 - s)}{\psi (w - a) + a} > 0, \quad \frac{\partial RIL}{\partial s} = -\frac{\psi w r(1 - s)(1 - \psi)}{\psi (w - a) + a} < 0, \]

\[ \frac{\partial RIL}{\partial \alpha} = \frac{\psi w r(1 - s)}{\psi (w - a) + a} > 0, \quad \frac{\partial RIL}{\partial \sigma} = \frac{\psi w r(1 - s) \left[ \ln(\alpha w) - \ln(1 - \alpha) \right]}{\psi w^\sigma + a} \left[ 1 - \frac{\psi w^\sigma}{\psi w^\sigma + a} \right] = 0, \]

Differentiations of equation (4.9'):

\[ RIL = \frac{\alpha wr(1 - s)}{\alpha (w - a) + a} \]
\[ \frac{\partial R_{IL}}{\partial w} = \frac{r(1-s)\alpha [1-\alpha]}{(\alpha(w-a) + a)^2} > 0, \quad \frac{\partial R_{IL}}{\partial r} = \frac{aw(1-s)}{(\alpha(w-a) + a)} > 0, \quad \frac{\partial R_{IL}}{\partial s} = -\frac{\alpha w}{(\alpha(w-a) + a)} < 0, \]
\[ \frac{\partial R_{IL}}{\partial a} = -\frac{\alpha w(1-s)[1-\alpha]}{(\alpha(w-a) + a)^2} > 0, \quad \frac{\partial R_{IL}}{\partial \alpha} = -\frac{wr(1-s)a}{(\alpha(w-a) + a)^2} > 0 \]

A.2

Differentiations of equation (4.14):

\[ AIL_{2,1} = \left[ r - \frac{w-a}{r w^\alpha + w} \right] \]
\[ \frac{\partial AIL}{\partial w} = \psi + \frac{\partial \psi}{\partial w} (w-a) - (1-r) = ? \]
\[ \frac{\partial AIL}{\partial r} = -w > 0, \quad \frac{\partial AIL}{\partial s} = 0, \quad \frac{\partial AIL}{\partial \alpha} = (1-\psi) > 0, \quad \frac{\partial AIL}{\partial \alpha} = \frac{(w-a)\psi(1-\psi)\sigma}{\alpha(1-\alpha)} > 0 \]
\[ \frac{\partial AIL}{\partial \sigma} = (w-a)\psi(1-\psi)[\ln(\alpha w) - \ln(1-\alpha)] = ? \]

Differentiations of equation (4.14'): \[ AIL_{2,1} = wr - (1-\alpha)(w-a), \]
\[ \frac{\partial AIL}{\partial w} = (r + \alpha - 1) = ?, \quad \frac{\partial AIL}{\partial r} = w > 0, \quad \frac{\partial AIL}{\partial s} = 0, \quad \frac{\partial AIL}{\partial \alpha} = (1-\alpha) > 0, \quad \frac{\partial AIL}{\partial \alpha} = (w-a) > 0 \]

Differentiations of equation (4.15):

\[ RIL_{2,1} = \frac{\psi(w-a) + a - w(1-r)}{\psi(w-a) + a} \]
\[ \frac{\partial RIL}{\partial w} = \frac{(1-r)\alpha(1-\psi) + \frac{1-\sigma}{\gamma w^\alpha + w} \left( w-a \right)^2(1-\psi) + aw(1-\psi) + w(\psi + r(a+w))}{\left( \psi(w-a) + a \right)^2} < 0 \]
\[ \frac{\partial RIL}{\partial r} = \frac{w}{\psi(w-a) + a} > 0, \quad \frac{\partial RIL}{\partial s} = 0, \quad \frac{\partial RIL}{\partial \alpha} = \frac{w(1-r)(1-\psi)}{\psi(w-a) + a} > 0, \]
\[ \frac{\partial RIL}{\partial \alpha} = \frac{w(1-r)(w-a)}{(\psi(w-a) + a)^3} \frac{\partial \psi}{\partial \alpha} = w(1-r)(w-a)\sigma \frac{\psi(1-\psi)}{\alpha(1-\alpha)} > 0, \]
\[ \frac{\partial RIL}{\partial \sigma} = \frac{w(1-r)(w-a)}{(\psi(w-a) + a)^2} \frac{\partial \psi}{\partial \sigma} = \frac{w(1-r)(w-a)\psi(1-\psi)[\ln(\alpha w) - \ln(1-\alpha)]}{(\psi(w-a) + a)^2} = ? \]

Differentiations of equation (4.15'): \[ RIL_{2,1} = \frac{\alpha(w-a) + a - w(1-r)}{\alpha(w-a) + a} \]
\[
\begin{align*}
\frac{\partial \mathcal{RIL}}{\partial w} = \frac{(1-r)(1-\alpha)a}{\alpha(w-a)+a} < 0, \quad \frac{\partial \mathcal{RIL}}{\partial r} = \frac{w}{\alpha(w-a)+a} > 0, \quad \frac{\partial \mathcal{RIL}}{\partial s} = 0, \\
\frac{\partial \mathcal{RIL}}{\partial \alpha} = \frac{w(1-r)(1-\alpha)}{(\alpha(w-a)+a)^2} > 0, \quad \frac{\partial \mathcal{RIL}}{\partial \alpha} = \frac{w(1-r)(w-a)}{(\alpha(w-a)+a)^2} > 0
\end{align*}
\]

A.3

Differentiations of equation (4.40):

\[
\begin{align*}
ACWL_1 &= \left( \frac{\alpha(yw^\sigma)^{1-\lambda} + 1 - \alpha}{\alpha(yw^\sigma)^{1-\lambda} + 1 - \alpha} \right)^{\frac{1}{\lambda}} \left( \frac{y}{yw^\sigma + w} \right) wr(1-s) \\
\frac{\partial ACWL}{\partial w} &= XY\gamma(1-s) \left[ 1 - \frac{\sigma yw^\sigma + \sigma}{yw^\sigma + w} + \frac{\sigma \alpha(yw^\sigma)^{1-\lambda}}{\alpha(yw^\sigma)^{1-\lambda} + 1 - \alpha} \right] \\
&= XY\gamma(1-s) \left[ 1 - \frac{(1-\alpha)\alpha(yw^\sigma)^{1-\lambda} - \alpha yw^\sigma}{(1-\alpha)\alpha(yw^\sigma)^{1-\lambda} + \alpha yw^\sigma} \right] > 0 \\
\frac{\partial ACWL}{\partial s} &= -XY\gamma < 0, \quad \frac{\partial ACWL}{\partial r} = XY\gamma(1-s) > 0 \\
\frac{\partial ACWL}{\partial \alpha} &= XY\gamma(1-s) \left[ \frac{(1-\alpha)(\sigma yw^\sigma - 1) + (1-\sigma)\alpha(yw^\sigma)^{1-\lambda}}{\alpha(yw^\sigma)^{1-\lambda} + 1 - \alpha} \right] > 0 \\
\frac{\partial ACWL}{\partial \sigma} &= XY\gamma(1-s) \sigma \left[ \frac{\alpha(yw^\sigma)^{1-\lambda} + \sigma(yw^\sigma)^{1-\lambda} - 1}{\alpha(yw^\sigma)^{1-\lambda} + 1 - \alpha} \right] = ? \\
\frac{\partial ACWL}{\partial \alpha} &= XY\gamma(1-s) \left[ \frac{\sigma yw^\sigma ln(w)(1-\alpha) + \sigma yw^\sigma ln\left(\frac{aw}{l-\alpha}\right)}{\sigma yw^\sigma + w} + \frac{\alpha(yw^\sigma)^{1-\lambda} ln\left(\frac{aw}{l-\alpha}\right)}{\alpha(yw^\sigma)^{1-\lambda} + 1 - \alpha} - \frac{\gamma(yw^\sigma)^{1-\lambda} ln\left(\frac{aw}{l-\alpha}\right)}{\alpha(yw^\sigma)^{1-\lambda} + 1 - \alpha} \right] = ?
\end{align*}
\]

Differentiations of equation (4.40'):

\[
ACWL_1 = \left( \frac{w}{w} \right)^{1-\alpha} wr(1-s)
\]
\[ \frac{\partial ACWL}{\partial w} = \alpha r(1-s) \left( \frac{w}{w} \right)^{1-\alpha} > 0, \quad \frac{\partial ACWL}{\partial s} = -\omega r \left( \frac{w}{w} \right)^{1-\alpha} < 0, \quad \frac{\partial ACWL}{\partial r} = \left( \frac{w}{w} \right)^{1-\alpha} w(1-s) > 0 \]

\[ \frac{\partial ACWL}{\partial w} = (1-\alpha) r (1-s) \left( \frac{w}{w} \right)^{-\alpha} > 0, \quad \frac{\partial ACWL}{\partial \alpha} = ACWL \left( \ln w - \ln \frac{w}{w} \right) = ? \]

Differentiations of equation (4.41):

\[ RCWL = \frac{wr(1-s)}{w-a} \]

\[ \frac{\partial RCWL}{\partial w} = -\frac{\alpha r(1-s)}{(w-a)^2} < 0 \]

\[ \frac{\partial RCWL}{\partial s} = -\frac{wr}{w-a} < 0 \]

\[ \frac{\partial RCWL}{\partial r} = \frac{w(1-s)}{w-a} > 0 \]

\[ \frac{\partial RCWL}{\partial \alpha} = \frac{wr(1-s)}{(w-a)^2} > 0 \]

\[ \frac{\partial ACWL}{\partial w} = \chi \left( \frac{w-a}{w\sigma} + w \right) \left( \frac{w-a}{w\sigma} \right)^{1-\alpha} - \left( \frac{\alpha(w(1-r)-a)^{-\lambda} + (1-\alpha)(sr)^{-\lambda}}{w-a} \right)^{\frac{1}{\lambda}} \]

where \( \chi = \frac{(w-a)(\bar{w}^\sigma + \bar{w})\alpha(\bar{w}^\sigma)^{1-\lambda} + (1-\alpha)^{\frac{1}{\lambda}}}{(\bar{w}^\sigma + \bar{w})\alpha(\bar{w}^\sigma)^{1-\lambda} + (1-\alpha)^{\frac{1}{\lambda}}} \]

\[ \frac{\partial ACWL}{\partial r} = \omega \left[ \alpha(w(1-r)-a)^{-\lambda} - (1-\alpha)(sr)^{-\lambda} \right] = ? \]

\[ \frac{\partial ACWL}{\partial s} = -\omega \left[ (1-\alpha) (sr)^{-\lambda} r \right] < 0 \]

where \( \omega = \frac{(\bar{w}^\sigma + \bar{w})\alpha(w(1-r)-a)^{1-\lambda} + (1-\alpha)(sr)^{1-\lambda}}{(\bar{w}^\sigma + \bar{w})\alpha(\bar{w}^\sigma)^{1-\lambda} + (1-\alpha)} \)
Differentiations of equation (4.45):

\[ ACWL = \left( \frac{w}{w} \right)^{1-a} \left( w - a - (w(1-r) - a)^{a} \right) \frac{(s_{r}w)^{1-a}}{\alpha^{a}(1-\alpha)^{1-a}} \]

\[
\frac{\partial ACWL}{\partial w} = \left( \frac{w}{w} \right)^{1-a} \left[ \alpha + \frac{a}{w} \left( 1 - a - \left( \frac{w(s_{r}w)^{1-a}}{(w(1-r) - a)(1-\alpha)} \right)^{1-a} \right)(1-r) \right] = ?
\]

\[
\frac{\partial ACWL}{\partial w} = (1-a) \left( \frac{w}{w} \right)^{1-a} \left[ 1 - \left( \frac{w(s_{r}w)^{1-a}}{(w(1-r) - a)(1-\alpha)} \right)^{1-a} \left( \frac{w(1-r) - a}{w} \right) \right] > 0
\]

\[
\frac{\partial ACWL}{\partial r} = \frac{w^{1-a} w^{-a} \left( s_{r} \right)^{1-a}}{(w(1-r) - a)(1-\alpha)^{1-a}} \left[ 1 - \left( 1 - \alpha \right) \left( w - a \right) \right] > 0
\]

\[
\frac{\partial ACWL}{\partial s} = \left( \frac{r^{1-a} w^{1-a} \left( w(1-r) - a \right)^{1-a}}{\alpha^{a}(1-\alpha)^{1-a}} \right) < 0
\]

\[
\frac{\partial ACWL}{\partial \alpha} = ACWL \left( \ln \left( \frac{w}{w} \right) + \left( \frac{w(1-r) - a}{w} \right)^{a} \left( \frac{s_{r}w^{a}}{\alpha^{a}(1-\alpha)^{1-a}} \right) - 1 \right) \ln \left( \frac{s_{r}w^{a}}{(w(1-r) - a)(1-\alpha)^{1-a}} \right) = ?
\]

Differentiations of equation (4.46):

\[ RCWL = 1 - \frac{w^{y} + w^{-a} \left( \alpha(w(1-r) - a)^{1-a} + (1-\alpha)(s_{r})^{1-a} \right)}{w^{y} + w^{-a} \left( \alpha(w(1-r) - a)^{1-a} + (1-\alpha)(s_{r})^{1-a} \right)} \]

\[
\frac{\partial RCWL}{\partial w} = -X Y \left[ \frac{\sigma w^{\alpha-1} + 1}{w^{\alpha} + w^{-a}} - \frac{\alpha(1-r)(w(1-r) - a)^{-1}}{\alpha(w(1-r) - a)^{1-a} + (1-\alpha)(s_{r})^{1-a}} - \frac{\sigma w^{\alpha-1}(w^{y})^{-1}}{(\alpha w^{y})^{-1} + 1-\alpha} \right] = ?
\]

Differentiations of equation (4.46'):

\[
RCWL = 1 - \frac{(w(1-r) - a)^{a} (s_{r})^{1-a}}{(w-a)\alpha^{a}(1-\alpha)^{1-a}}
\]

\[
\frac{\partial RCWL}{\partial w} = \frac{(s_{r})^{1-a}}{\alpha^{a}(1-\alpha)^{1-a} (w-a)^{2} (w(1-r) - a)^{1-a}} \left( 1 - \frac{r + a}{w} \right) < 0
\]

\[
\frac{\partial RCWL}{\partial \alpha} = \frac{(s_{r})^{1-a} (w(1-r) - a)^{1-a}}{\alpha^{a}(1-\alpha)^{1-a} (w-a)^{2} (w(1-r) - a)^{1-a}} \left( 1 - \frac{r + a}{w} \right) < 0
\]

\[
\frac{\partial RCWL}{\partial r} = \frac{(w(1-r) - a)^{1-a}}{\alpha^{a}(1-\alpha)^{1-a} (w-a)^{2} (w(1-r) - a)^{1-a}} \left( 1 - \frac{r + a}{w} \right) > 0
\]

\[
\frac{\partial RCWL}{\partial s} = \frac{(w(1-r) - a)^{1-a}}{\alpha^{a}(1-\alpha)^{1-a} (w-a)^{2} (w(1-r) - a)^{1-a}} \left( 1 - \frac{r + a}{w} \right) > 0
\]

\[
\frac{\partial RCWL}{\partial \alpha} = \frac{(w(1-r) - a)^{1-a}}{\alpha^{a}(1-\alpha)^{1-a} (w-a)^{2} (w(1-r) - a)^{1-a}} \left( 1 - \frac{r + a}{w} \right) > 0
\]

\[ 288 \]
Appendix 5.A: Differentiation of Dose-Response Functions.

The dose-response coefficients for the entire illness event ($\mathcal{A}/\mathcal{P}$) and for the probability of occurrence ($\mathbb{q}/\mathcal{P}$) can be found by differentiating expressions (12a) and (13) with respect to pollution:

$$\frac{dH}{dP} = \left[ \frac{\partial q}{\partial E} + \frac{\partial q}{\partial A} \right] \times S + \left[ \frac{\partial S}{\partial M} \frac{\partial M}{\partial P} + \frac{\partial S}{\partial A} \frac{\partial A}{\partial P} \right] \times q$$

$$\frac{dq}{dP} = \frac{\partial q}{\partial E} \left( \frac{\partial E}{\partial P} + \frac{\partial E}{\partial A} \right) = \frac{\partial q}{\partial P} + \frac{\partial q}{\partial A}$$

As for the effects of income on these two alternative dose-response coefficients, they are given by:

$$\frac{d^2H}{dPdY} = \frac{\partial q}{\partial E} \left[ \frac{\partial E}{\partial A} \frac{\partial A}{\partial Y} \right] \times S$$

$$+ \left[ \frac{\partial q}{\partial E} \left( \frac{\partial E}{\partial P} + \frac{\partial E}{\partial A} \right) \right] \left[ \frac{\partial S}{\partial M} \frac{\partial M}{\partial Y} + \frac{\partial S}{\partial A} \frac{\partial A}{\partial Y} \right]$$

$$+ \left[ \frac{\partial S}{\partial M} \frac{\partial^2 M}{\partial M^2} \frac{\partial M}{\partial Y} + \frac{\partial S}{\partial A} \frac{\partial^2 A}{\partial A^2} \frac{\partial A}{\partial Y} \right] \times q$$

$$+ \left[ \frac{\partial S}{\partial M} \frac{\partial^2 M}{\partial M \partial P} + \frac{\partial S}{\partial A} \frac{\partial^2 A}{\partial A \partial P} \right] \times \frac{\partial q}{\partial E} \frac{\partial E}{\partial A}$$

$$\frac{dq}{dPdY} = \frac{\partial q}{\partial E} \left[ \frac{\partial E}{\partial A} \frac{\partial A}{\partial Y} + \frac{\partial E}{\partial Y} \frac{\partial A}{\partial P} + \frac{\partial E}{\partial A} \frac{\partial A}{\partial P} \right]$$

289
Appendix 5.B: Hypothetical Examples.

Hypothetical example 1: The effect of transferring a dose-response estimate from a rich country to a developing country.

In Table 1 below we have two cities, “rich” and “poor”, each with a population (POP) of 10 million. The “rich” city has 8 million rich inhabitants and 2 million poor ones, whereas the “poor” city has the opposite (2 million rich, 8 million poor). Furthermore, the number of Restricted Activity Days (RAD) due to the adverse health effect from a certain change in the amount of a particular air pollutant is 4 for the rich part of the population and 6 for the poor. From the calculations of the dose-response relationship in the last column on the right in Table 1, we see that the estimated slope in the rich city is 4.4 (i.e. \((8 \times 4 + 2 \times 6)/10\)). Transferring this estimate to the poor city where the underlying income distribution is different will lead to incorrect measurements, as the correct average effect there is \(b_{ij} = 5.6\) (i.e. \((2 \times 4 + 8 \times 6)/10\)).

<table>
<thead>
<tr>
<th>COUNTRY</th>
<th>SCENARIO</th>
<th>POP (rich)</th>
<th>POP (poor)</th>
<th>RAD (rich)</th>
<th>RAD (poor)</th>
<th>(b_{ij})</th>
</tr>
</thead>
<tbody>
<tr>
<td>RICH</td>
<td>Correct</td>
<td>8</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>4.4</td>
</tr>
<tr>
<td>POOR</td>
<td>Correct</td>
<td>2</td>
<td>8</td>
<td>4</td>
<td>6</td>
<td>5.6</td>
</tr>
</tbody>
</table>

Hypothetical example 2: The effect of using average WTP and dose-response measures in calculating costs when both vary with income.

The only difference between scenarios A and B in Table 2 is the estimated amount of restricted activity days (RAD) in the rich and the poor population (columns 7 and 8, respectively) resulting from an increase in air pollution - the average effect in the population as a whole (column 9) is the same. In scenario A only the average RAD estimate is available and this is used together with the WTP to calculate the illness cost. Since the WTPs are known to be different for the poor and the rich in scenario A, using the average RAD to calculate illness cost means that we are implicitly assuming the health effect to be equal for the poor and the rich. The difference between the two scenarios suffices to give two different values to the health change in the population, i.e. different illness costs, as observable in the last column on the right hand side. Scenario B and C only differ in the differing WTP of the poor and the rich – the average effect is the same and the different health effects on the two groups is here accounted for. In scenario
C the different RAD-estimates for the poor and the rich part of the population are available, however, only the average WTP is known. Using the average WTP to calculate the costs therefore means that we are implicitly assuming poor and rich people to have the same WTP to avoid illness events. Here again, the difference between the two scenarios leads to different quantifications of the adverse health impact due to air pollution. In this particular example we therefore see that the two scenarios that rely partly on averaged information, i.e. scenarios A and C, exaggerate the illness cost of air pollution compared to the true scenario, i.e. scenario B.

Table 2: The effect on the estimated cost of relying on averages of WTP and RAD.

<table>
<thead>
<tr>
<th>SCENARIO</th>
<th>POP (rich)</th>
<th>POP (poor)</th>
<th>WTP (rich)</th>
<th>WTP (poor)</th>
<th>WTP</th>
<th>RAD (rich)</th>
<th>RAD (poor)</th>
<th>RAD $b_j$</th>
<th>Illness Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>3</td>
<td>4.5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>225</td>
</tr>
<tr>
<td>B (correct)</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>3</td>
<td>4.5</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>210</td>
</tr>
<tr>
<td>C</td>
<td>5</td>
<td>5</td>
<td>4.5</td>
<td>4.5</td>
<td>4.5</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>225</td>
</tr>
</tbody>
</table>

Hypothetical example 3: The implication on valuation of not capturing the health effect on the lower income groups with the health measurement.

The correct scenarios given in table 3, i.e. rows 2 and 5, are the same as in table 1. However, instead of assuming the health effect to be the same for the poor as for the rich when a measurement of the health effect on the poor part of the population is missing, the health effect measurement used in this example does not register the morbidity experienced by the poor. In this case the downward bias in the estimated slope of the dose-response function is even larger, with the true average being 4.4 in the rich country, and 5.6 in the poor country, and the measured slope being 3.2 (i.e. $(8\times4)/10$) and 0.8 (i.e. $(2\times4)/10$), respectively. As for the quantification of the adverse health effects, not capturing the health effect on a part of the population will clearly lead to wrong results. In the case of the "rich" country we can see from the last column that the true cost estimate is 228, whereas using the rich individual’s WTP or the average WTP and not capturing the health effect on the poor yields estimates of 192 and 173, respectively.

Table 3: Bias in the measurement of morbidity (information bias).

<table>
<thead>
<tr>
<th>COUNTRY</th>
<th>SCENARIO</th>
<th>POP (rich)</th>
<th>POP (poor)</th>
<th>RAD (rich)</th>
<th>RAD (poor)</th>
<th>RAD $b_j$</th>
<th>WTP (rich)</th>
<th>WTP (poor)</th>
<th>WTP</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>RICH</td>
<td>Correct</td>
<td>8</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>4.4</td>
<td>6</td>
<td>3</td>
<td>5.4</td>
<td>228</td>
</tr>
<tr>
<td>RICH</td>
<td>False</td>
<td>8</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>3.2</td>
<td>6</td>
<td>3</td>
<td>5.4</td>
<td>192</td>
</tr>
<tr>
<td>RICH</td>
<td>False</td>
<td>8</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>3.2</td>
<td>6</td>
<td>3</td>
<td>5.4</td>
<td>173</td>
</tr>
<tr>
<td>POOR</td>
<td>Correct</td>
<td>2</td>
<td>8</td>
<td>4</td>
<td>6</td>
<td>5.6</td>
<td>6</td>
<td>3</td>
<td>5.4</td>
<td>192</td>
</tr>
<tr>
<td>POOR</td>
<td>False</td>
<td>2</td>
<td>8</td>
<td>4</td>
<td>0</td>
<td>0.8</td>
<td>6</td>
<td>3</td>
<td>5.4</td>
<td>43</td>
</tr>
</tbody>
</table>

291

Dr. David Maddison, Senior Research Fellow at the Centre for Social and Economic Research on the Global Environment (CSERGE), took the initiative for undertaking a meta-analysis of air pollution mortality studies. In particular, his idea was to include moderator variables that potentially influence the slope of the dose-response function. The prime purpose of the paper was to quantify and value the health impacts of air pollution in densely populated cities of the developing world.

In the Maddison and Gaarder (MG) study we analysed a smaller sample (13 observations) than in the present paper (70 observations), using slightly different selection rules for the sample. In the present paper I furthermore argue that one of the observations in that sample should have been excluded. As for the selection of moderator variables, Dr. Maddison suggested including the level of ambient particle concentration and the percentage of the population over 65 years of age, whereas I suggested including income and income inequality. In the present paper, I additionally include baseline health, education, and health services as moderator variables in my analysis, and update all of the data used to capture the moderator variables. In the MG-study we used the method of variance-weighted least squares (VWLS), thereby implicitly assuming that all the variance among the study effects other than sampling variance can be explained as a function of the study characteristics we chose to include. In the present study I argue that that assumption is unrealistic, and propose to use Empirical Bayes method to take into account unexplained inter-study variability.

I show in this paper how the data, methodology, and sample corrections significantly change the findings from those presented in the MG-paper. Then I proceed by analysing the larger sample gathered specifically for the current paper, and discussing the findings.

In addition to the above differences, the present paper reviews the existing meta-analysis literature, carries out a thorough sensitivity-analysis, and discusses in detail the weaknesses attached to such a study.
Appendix 6.B: Mathematical Derivation of the Equations in Section 6.4

This appendix is following closely the derivation in Raudenbush and Bryk (1985). The maximum likelihood estimators $\delta_j$ and $\gamma$, with known (or estimated) $\nu_i$ and $\tau_i$ can be derived by the maximum likelihood method:

$$L(\delta, \gamma; d) = \prod_i f_i(d_i|\delta_i) g(\delta_i|\nu_i, \gamma)$$

$$= \left[ \prod_i (2\pi\nu_i)^{-1/2} \exp\left\{-\frac{1}{2} \sum (d_i - \delta_i)^2 / \nu_i \right\} \right] \times \left[ \prod_i (2\pi \tau_i)^{-1/2} \exp\left\{-\frac{1}{2} \sum (d_i - W_i' \gamma)^2 / \tau_i \right\} \right]$$

Taking the derivatives of $\log L$ with respect to $\delta_i$ and setting it equal to 0 yields:

$$(d_i - \delta_i)/\nu_i = (\delta_i - W_i' \gamma)/\tau_i$$

Solving this equation for $\delta_i$ we obtain

$$\delta_i = \nu_i (\nu_i + \tau_i)^{-1} W_i' d_i + (\nu_i + \tau_i)^{-1} W_i' \gamma$$

Setting $\partial \log L / \partial \gamma = 0$ yields

$$\sum \delta_i = \sum W_i \gamma W_i' = \gamma \sum W_i W_i'$$

If we transpose both sides and substitute for $\delta_i$ from equation (FC) we obtain the following expression for $\gamma$:

$$\gamma = \left( \sum \lambda_i W_i W_i' \right)^{-1} \sum \lambda_i W_i d_i$$

The approximate likelihood of the $d_i$ is given by the marginal density of the vector $d = (d_1, d_2, \ldots, d_k)$:

$$L(\gamma, \tau^2; d) = \left\{ \prod_i [2\pi(v_i + \tau^2)]^{-1/2} \right\} \exp\left\{ -\frac{1}{2} \sum (v_i + \tau^2)^{-1} (d_i - W_i' \gamma)^2 \right\}$$

$$= \left\{ \prod_i [2\pi(v_i + \tau^2)]^{-1/2} \right\} \exp\left\{ -\frac{1}{2} Q \right\}$$

where

$$Q = \sum (v_i + \tau^2)^{-1} (d_i - W_i' \gamma)^2 = \sum (v_i + \tau^2)^{-1} (d_i - W_i' \gamma *)^2 + (\gamma ^* - \gamma) \sum (v_i + \tau^2)^{-1} W_i W_i' (\gamma ^* - \gamma)$$

$$= Q_1 + Q_2$$

Under the distributional assumption of equation (6.10) we find that $\gamma ^*$ is multivariate normal with mean vector $\gamma$ and variance $\left[ \sum (v_i + \tau^2)^{-1} W_i W_i' \right]^{-1}$.

The density of $\gamma ^*$ is hence;
\[ f(\gamma^*, \tau^2, \gamma) = (2\pi)^{-n/2} \left| \sum (v_i + \tau^2)^{-1} W_i \right|^{1/2} \exp\left\{ -\frac{1}{2} \Omega_2 \right\} \]

The density of \( d \) conditional on \( \gamma^* \) can now be expressed as

\[
g(d|\gamma^*, \tau^2) = \frac{L(\gamma^*, \tau^2; d)}{f(\gamma^*, \tau^2, \gamma)} = \frac{\left\{ \prod_i [2\pi (v_i + \tau^2)]^{-1/2} \right\} \exp\left\{ -\frac{1}{2} (\Omega_1 + \Omega_2) \right\}}{(2\pi)^{-n/2} \left| \sum (v_i + \tau^2)^{-1} W_i \right|^{1/2} \exp\left\{ -\frac{1}{2} \Omega_2 \right\}}
\]

\[
= (2\pi)^{-1/2(n-\theta)} \prod_i (v_i + \tau^2)^{-1/2} \sum (v_i + \tau^2)^{-1} W_i W_i^{-1/2} \exp\left\{ -\frac{1}{2} \Omega_1 \right\}
\]

It follows from this expression that the log of the likelihood is proportional to expression (6.10).
Appendix 6.C: Do-file in STATA to Maximise the Likelihood Function

The main difference between the two methods presented below is that method If requires that the likelihood function meets the linear-form restrictions, whereas method d0 does not. Since our likelihood function does meet the linear-form restrictions, we can use either method. Note that the specifications written in italics were altered according to the exact model we wanted to run, and that the line-numbering does not belong to the do-file but was added in order to simplify this discussion. The program name in lines 1 and 14 (1 and 10 in method If) must be changed whenever a change is made to the do-file. In addition, the variables entered into lines 8 and 14 of the do-file may be altered (lines 5 and 10 in method If), as well as the sample size specified in line 11 (line 8 in method If).

The parenthesis at the end of line 14 (line 10 for method If) indicates that if we want to restrict our sample, this can be done by adding `if` and then the relevant restriction (the parenthesis is there to indicate that this is a choice feature – however, if the feature is used the parenthesis should be removed). Note that when restrictions are made on the sample, the sample size in line 11 (line 8 in method If) must be adjusted to reflect this. Finally, it is optional to add 'difficult' to the maximize-command in line 15 (line 11 in method If).

**Method d0:**
1. program define d0prog
2. version 6
3. args todo b lnf
4. tempvar t1 t2 detx Indetx
5. mleval `t1' = `b'
6. mleval `t2' = `b', eq(2)
7. quietly {
8. matrix accum X = poll over65 gnp dist educ dale [iweight=(se^2+`t2')^(-1)]
9. gen double `detx' = det(X)
10. gen double `Indetx' = ln(`detx')
11. mlsum `lnf' = -ln(se^2+`t2')-`Indetx'/70-((se^2+`t2')^(-1)*($ML_y1-t1')^2)
12. }
13. end
14. ml model d0 d0prog (regco = poll over65 gnp dist educ dale) / tau2 (if restriction)
15. ml maximize(, difficult)
Method If:
1. program define Ifprog
2. version 6
3. args lnf tl t2
4. tempvar detx Indetx
5. matrix accum X = poll over65 gnp dist educ dale [iweight=( se^2+`t2')^(-1)]
6. quietly gen double `detx' = det(X)
7. quietly gen double `Indetx' = ln(`detx')
8. quietly replace `lnf' = -ln(se^2+`t2')-lnindetx'/70-((se^2+`t2')^(-1)*($ML_y1-`t1')^2)
9. end
10. ml model If Ifprog (regco = poll over65 gnp dist educ dale) / tau2 (if restriction)
11. ml maximize(, difficult)
Table 6.11: Data based on information given in studies

<table>
<thead>
<tr>
<th>OBS.</th>
<th>REFERENCE</th>
<th>PUBL. YEAR</th>
<th>COUNTRY</th>
<th>CITY</th>
<th>PERIOD</th>
<th>REGCO (Regression coefficient)</th>
<th>SE (Standard error)</th>
<th>Change in daily mortality(%)/10μg/m³ increase in PM₁₀</th>
<th>POLL (average PM₁₀ level for period)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Touloumi et al.</td>
<td>1996</td>
<td>Greece</td>
<td>Athens</td>
<td>1987-1991</td>
<td>0.000480</td>
<td>0.000110</td>
<td>0.48</td>
<td>84.4</td>
</tr>
<tr>
<td>2</td>
<td>Sunyer et al.</td>
<td>1996</td>
<td>Spain</td>
<td>Barcelona</td>
<td>1985-1991</td>
<td>0.000677</td>
<td>0.000198</td>
<td>0.68</td>
<td>42.4</td>
</tr>
<tr>
<td>3</td>
<td>Bacharova et al.</td>
<td>1996</td>
<td>Slovak Rep.</td>
<td>Bratislava</td>
<td>1987-1991</td>
<td>0.000022</td>
<td>0.000036</td>
<td>0.02</td>
<td>49.2</td>
</tr>
<tr>
<td>4</td>
<td>Wojtyniak et al.</td>
<td>1996</td>
<td>Poland</td>
<td>Cracow</td>
<td>1977-1989</td>
<td>0.000173</td>
<td>0.000083</td>
<td>0.17</td>
<td>73.3</td>
</tr>
<tr>
<td>5</td>
<td>Wojtyniak et al.</td>
<td>1996</td>
<td>Poland</td>
<td>Lodz</td>
<td>1977-1990</td>
<td>0.000213</td>
<td>0.000085</td>
<td>0.21</td>
<td>57.3</td>
</tr>
<tr>
<td>6</td>
<td>Wojtyniak et al.</td>
<td>1996</td>
<td>Poland</td>
<td>Poznan</td>
<td>1983-1990</td>
<td>0.000175</td>
<td>0.000159</td>
<td>0.18</td>
<td>34.0</td>
</tr>
<tr>
<td>7</td>
<td>Wojtyniak et al.</td>
<td>1996</td>
<td>Poland</td>
<td>Wroclaw</td>
<td>1979-1989</td>
<td>0.000075</td>
<td>0.000129</td>
<td>0.08</td>
<td>54.3</td>
</tr>
<tr>
<td>8</td>
<td>Anderson et al.</td>
<td>1996</td>
<td>UK</td>
<td>London</td>
<td>1987-1992</td>
<td>0.001204</td>
<td>0.000315</td>
<td>1.21</td>
<td>14.6</td>
</tr>
<tr>
<td>9</td>
<td>Spix and Wichmann</td>
<td>1996</td>
<td>Germany</td>
<td>Koeln</td>
<td>1975-1985</td>
<td>0.000267</td>
<td>0.000239</td>
<td>0.27</td>
<td>37.4</td>
</tr>
<tr>
<td>10</td>
<td>Zmirou et al.</td>
<td>1996</td>
<td>France</td>
<td>Lyon</td>
<td>1985-1990</td>
<td>0.000258</td>
<td>0.000525</td>
<td>0.26</td>
<td>29.3</td>
</tr>
<tr>
<td>11</td>
<td>Cropper et al.</td>
<td>1997</td>
<td>India</td>
<td>New Delhi</td>
<td>1991-1994</td>
<td>0.000413</td>
<td>0.000182</td>
<td>0.41</td>
<td>207.9</td>
</tr>
<tr>
<td>12</td>
<td>Ostro et al.</td>
<td>1995</td>
<td>Chile</td>
<td>Santiago</td>
<td>1989-1991</td>
<td>0.000750</td>
<td>0.000130</td>
<td>0.75</td>
<td>115.4</td>
</tr>
<tr>
<td>13</td>
<td>Perez et al.</td>
<td>1999</td>
<td>Spain</td>
<td>Cartagena</td>
<td>1992-1996</td>
<td>-0.000164</td>
<td>0.000093</td>
<td>-0.16</td>
<td>31.6</td>
</tr>
<tr>
<td>14</td>
<td>Contin et al.</td>
<td>1999</td>
<td>Spain</td>
<td>Bilbao</td>
<td>1992-1996</td>
<td>0.001503</td>
<td>0.000488</td>
<td>1.51</td>
<td>43.1</td>
</tr>
<tr>
<td>15</td>
<td>Burillo et al.</td>
<td>1999</td>
<td>Spain</td>
<td>Valencia</td>
<td>1994-1996</td>
<td>0.001262</td>
<td>0.000494</td>
<td>1.27</td>
<td>44.2</td>
</tr>
<tr>
<td>16</td>
<td>Labaca et al.</td>
<td>1999</td>
<td>Spain</td>
<td>Madrid</td>
<td>1992-1995</td>
<td>-0.000441</td>
<td>0.000225</td>
<td>-0.44</td>
<td>37.8</td>
</tr>
<tr>
<td>17</td>
<td>Trunk et al.</td>
<td>1999</td>
<td>Spain</td>
<td>Vigo</td>
<td>1991-1994</td>
<td>0.000349</td>
<td>0.000330</td>
<td>0.35</td>
<td>98.1</td>
</tr>
<tr>
<td>18</td>
<td>Daponte-Codina et al.</td>
<td>1999</td>
<td>Spain</td>
<td>Huelva</td>
<td>1993-1996</td>
<td>0.002460</td>
<td>0.001361</td>
<td>2.49</td>
<td>42.5</td>
</tr>
<tr>
<td>19</td>
<td>Martinez et al.</td>
<td>1999</td>
<td>Spain</td>
<td>Gijon</td>
<td>1993-1996</td>
<td>0.001088</td>
<td>0.000696</td>
<td>1.09</td>
<td>45.6</td>
</tr>
<tr>
<td>20</td>
<td>Martinez et al.</td>
<td>1999</td>
<td>Spain</td>
<td>Oviedo</td>
<td>1993-1997</td>
<td>-0.001149</td>
<td>0.000961</td>
<td>-1.14</td>
<td>43.5</td>
</tr>
<tr>
<td>21</td>
<td>Boilos et al.</td>
<td>1999</td>
<td>Spain</td>
<td>Vitoria-Gasteiz</td>
<td>1990-1994</td>
<td>0.000628</td>
<td>0.000487</td>
<td>0.63</td>
<td>51.2</td>
</tr>
<tr>
<td>22</td>
<td>Blasco et al.</td>
<td>1999</td>
<td>Spain</td>
<td>Castellon</td>
<td>1991-1995</td>
<td>0.001499</td>
<td>0.001020</td>
<td>1.51</td>
<td>24.6</td>
</tr>
<tr>
<td>23</td>
<td>Ontoso et al.</td>
<td>1999</td>
<td>Spain</td>
<td>Pamplona</td>
<td>1991-1995</td>
<td>0.002941</td>
<td>0.002470</td>
<td>2.98</td>
<td>21.7</td>
</tr>
<tr>
<td>24</td>
<td>Ocana-Riola et al.</td>
<td>1999</td>
<td>Spain</td>
<td>Sevilla</td>
<td>1992-1996</td>
<td>-0.002013</td>
<td>0.000650</td>
<td>-1.99</td>
<td>45.1</td>
</tr>
<tr>
<td></td>
<td>Author(s)</td>
<td>Year</td>
<td>Location</td>
<td>Start-End</td>
<td>Mortality Rate 1</td>
<td>Mortality Rate 2</td>
<td>Mortality Rate 3</td>
<td>Mortality Rate 4</td>
<td>Mortality Rate 5</td>
</tr>
<tr>
<td>---</td>
<td>------------------------------</td>
<td>------</td>
<td>-------------------</td>
<td>-----------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>25</td>
<td>Ballester et al.</td>
<td>1996</td>
<td>Spain Valencia</td>
<td>1991-1993</td>
<td>0.000895</td>
<td>0.000310</td>
<td>0.90</td>
<td>67.7</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Schwartz</td>
<td>1991</td>
<td>USA Cincinnati</td>
<td>1977-1982</td>
<td>0.001059</td>
<td>0.000305</td>
<td>1.07</td>
<td>41.8</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Kinney et al.</td>
<td>1995</td>
<td>USA LA</td>
<td>1985-1990</td>
<td>0.000488</td>
<td>0.000266</td>
<td>0.49</td>
<td>58.0</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Schwartz</td>
<td>1993</td>
<td>USA Birmingham, AL</td>
<td>1985-1988</td>
<td>0.001044</td>
<td>0.000415</td>
<td>1.05</td>
<td>47.9</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>Schwartz and Dockery</td>
<td>1992</td>
<td>USA Steubenville</td>
<td>1974-1984</td>
<td>0.000693</td>
<td>0.000149</td>
<td>0.70</td>
<td>61.1</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Ito et al.</td>
<td>1995</td>
<td>USA Cook c.</td>
<td>1985-1990</td>
<td>0.000583</td>
<td>0.000218</td>
<td>0.58</td>
<td>38.0</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>Schwartz and Dockery</td>
<td>1992</td>
<td>USA Philadelphia</td>
<td>1973-1980</td>
<td>0.001202</td>
<td>0.000238</td>
<td>1.21</td>
<td>42.5</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>Pope et al.</td>
<td>1992</td>
<td>USA Utah valley</td>
<td>1985-1989</td>
<td>0.001470</td>
<td>0.000310</td>
<td>1.48</td>
<td>47.0</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>Pope</td>
<td>1999</td>
<td>USA Ogden</td>
<td>1985-1995</td>
<td>0.001360</td>
<td>0.000620</td>
<td>1.37</td>
<td>32.1</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>Pope</td>
<td>1999</td>
<td>USA Salt Lake City</td>
<td>1985-1995</td>
<td>0.000460</td>
<td>0.000230</td>
<td>0.46</td>
<td>41.2</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>Pope</td>
<td>1999</td>
<td>USA Provo/Orem</td>
<td>1985-1995</td>
<td>0.000870</td>
<td>0.000270</td>
<td>0.87</td>
<td>38.4</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>Dockery et al.</td>
<td>1992</td>
<td>USA St. Louis</td>
<td>1985-1986</td>
<td>0.001500</td>
<td>0.000690</td>
<td>1.51</td>
<td>27.6</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>Dockery et al.</td>
<td>1992</td>
<td>USA Kingston</td>
<td>1985-1986</td>
<td>0.001600</td>
<td>0.000149</td>
<td>1.61</td>
<td>30.0</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>Schwartz</td>
<td>1994</td>
<td>USA Detroit</td>
<td>1973-1982</td>
<td>0.000993</td>
<td>0.000264</td>
<td>1.00</td>
<td>47.9</td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>Kelsall</td>
<td>1997</td>
<td>USA Philadelphia</td>
<td>1974-1988</td>
<td>0.000603</td>
<td>0.000194</td>
<td>0.60</td>
<td>37.0</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>Mar et al.</td>
<td>2000</td>
<td>USA Phoenix</td>
<td>1995-1997</td>
<td>0.001060</td>
<td>0.000535</td>
<td>1.08</td>
<td>46.5</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>Schwartz et al.</td>
<td>1996</td>
<td>USA Boston</td>
<td>1979-1986</td>
<td>0.001193</td>
<td>0.000252</td>
<td>1.21</td>
<td>24.5</td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>Schwartz et al.</td>
<td>1996</td>
<td>USA Knoxville</td>
<td>1980-1987</td>
<td>0.000896</td>
<td>0.000430</td>
<td>0.91</td>
<td>32.0</td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>Schwartz et al.</td>
<td>1996</td>
<td>USA St. Louis</td>
<td>1979-1987</td>
<td>0.000598</td>
<td>0.000228</td>
<td>0.61</td>
<td>30.6</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>Schwartz et al.</td>
<td>1996</td>
<td>USA Steubenville</td>
<td>1979-1987</td>
<td>0.000896</td>
<td>0.000379</td>
<td>0.91</td>
<td>45.6</td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>Schwartz et al.</td>
<td>1996</td>
<td>USA Portage</td>
<td>1979-1987</td>
<td>0.000698</td>
<td>0.000532</td>
<td>0.71</td>
<td>17.8</td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>Schwartz et al.</td>
<td>1996</td>
<td>USA Topeka</td>
<td>1979-1988</td>
<td>-0.000501</td>
<td>0.000276</td>
<td>-0.51</td>
<td>26.7</td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>Ostro et al.</td>
<td>1999</td>
<td>USA Coachella</td>
<td>1989-1992</td>
<td>0.000900</td>
<td>0.000400</td>
<td>0.91</td>
<td>62.0</td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>Hong et al.</td>
<td>1999</td>
<td>South Korea Incheon</td>
<td>1995-1996</td>
<td>0.000700</td>
<td>0.000300</td>
<td>0.70</td>
<td>71.2</td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>Lee et al.</td>
<td>1999</td>
<td>South Korea Seoul</td>
<td>1991-1995</td>
<td>0.000904</td>
<td>0.000181</td>
<td>0.91</td>
<td>50.9</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>Lee et al.</td>
<td>1999</td>
<td>South Korea Ulsan</td>
<td>1991-1995</td>
<td>-0.000018</td>
<td>0.000362</td>
<td>-0.02</td>
<td>39.7</td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>Peters et al.</td>
<td>2000</td>
<td>Czech Coal Basin</td>
<td>1982-1994</td>
<td>0.000678</td>
<td>0.000273</td>
<td>0.68</td>
<td>66.7</td>
<td></td>
</tr>
<tr>
<td>52</td>
<td>Peters et al.</td>
<td>2000</td>
<td>Germany Bavaria</td>
<td>1982-1994</td>
<td>0.000073</td>
<td>0.000250</td>
<td>0.07</td>
<td>28.4</td>
<td></td>
</tr>
<tr>
<td>53</td>
<td>Xu et al.</td>
<td>2000</td>
<td>China Shenyang</td>
<td>1992</td>
<td>0.000315</td>
<td>0.000100</td>
<td>0.32</td>
<td>236.5</td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>Xu et al.</td>
<td>1994</td>
<td>China Beijing</td>
<td>1989</td>
<td>0.000200</td>
<td>0.000164</td>
<td>0.20</td>
<td>206.3</td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>Borja-Aburto et al.</td>
<td>1997</td>
<td>Mexico Mexico city</td>
<td>1990-1992</td>
<td>0.000887</td>
<td>0.000164</td>
<td>0.89</td>
<td>118.8</td>
<td></td>
</tr>
<tr>
<td>56</td>
<td>Castillejos et al.</td>
<td>2000</td>
<td>Mexico Mexico city</td>
<td>1992-1995</td>
<td>0.001813</td>
<td>0.000426</td>
<td>1.83</td>
<td>44.6</td>
<td></td>
</tr>
<tr>
<td>57</td>
<td>Verhoeff et al.</td>
<td>1996</td>
<td>Netherlands Amsterdam</td>
<td>1986-1992</td>
<td>0.000602</td>
<td>0.000379</td>
<td>0.60</td>
<td>38.0</td>
<td></td>
</tr>
<tr>
<td>58</td>
<td>Hoek et al.</td>
<td>1997</td>
<td>Netherlands Rotterdam</td>
<td>1983-1991</td>
<td>0.000975</td>
<td>0.000389</td>
<td>0.98</td>
<td>23.1</td>
<td></td>
</tr>
</tbody>
</table>
Sources: The data are taken from the studies referenced in column 1 of the table.

Notes:
1. The information concerning the effect from exposure to air pollution on the risk of mortality uncovered by the various studies was expressed in a number of alternative ways. For those studies not expressing the effect in terms of the original regression coefficient from Poisson model, the results had to be converted. The natural logarithm of relative risk (RR) divided by number of units of air pollution increase, i.e. $\Delta P$, gives the original regression coefficient, $\beta$ (relative risk indicates the ratio of the probability of occurrence of a given effect between two different exposure levels or exposure groups): i.e. $\beta = \ln(RR) / \Delta P$. If the effect is given in terms of percentage increase in mortality associated with a certain increase in pollutant level, the relative risk may be obtained by dividing the percentage increase by 100 and adding one. Finally, when the effect is given in terms of a coefficient of elasticity (i.e. change in mortality rate due to a change in the logarithm of pollution), the relative risk is obtained by multiplying the coefficient of elasticity, $\varepsilon_{M,P}$, with the change in logarithm of pollution and taking its exponential: i.e. $RR = \exp(\varepsilon_{M,P} \ast \Delta \ln(P))$.
2. The standard error was expressed in a manner consistent with the estimated effect, and was hence converted in a similar manner (see footnote 1). Some studies only gave information on the confidence interval, and the standard error was calculated from this information.
3. The change in daily mortality (%) per 10$\mu$g/m$^3$ increase in PM$_{10}$ was calculated as follows: $(\text{EXP}(\beta \ast 10) - 1) \ast 100$, where $\beta$ is the original regression coefficient.
4. Each study in the meta-analysis supplied mean values of daily data over the study period for either TSP, BS, or PM. TSP and PM$_{10}$ were converted to PM$_{10}$ using the factors of 0.55 and 0.77, respectively, and black smoke was considered equal to PM$_{10}$. Note that this implied dividing the estimated coefficients in studies using the TSP and PM$_{10}$ measures by 0.55 and 0.77, respectively, in order to convert these into being PM$_{10}$ or BS effects.
Table 6. 12: Data gathered from SIMA and WHO

<table>
<thead>
<tr>
<th>OBS</th>
<th>OVER65</th>
<th>GNP</th>
<th>GINI</th>
<th>DIST</th>
<th>EDUC</th>
<th>DALE</th>
<th>LEAB</th>
<th>PHYSICIANS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>13.68</td>
<td>10815</td>
<td>32.70</td>
<td>353663</td>
<td>82.84</td>
<td>72.50</td>
<td>76.79</td>
<td>3.40</td>
</tr>
<tr>
<td>2</td>
<td>13.08</td>
<td>10727</td>
<td>32.50</td>
<td>348628</td>
<td>90.90</td>
<td>72.80</td>
<td>76.44</td>
<td>3.77</td>
</tr>
<tr>
<td>3</td>
<td>9.99</td>
<td>8313</td>
<td>19.50</td>
<td>162099</td>
<td>89.74</td>
<td>66.60</td>
<td>71.03</td>
<td>3.30</td>
</tr>
<tr>
<td>4</td>
<td>9.78</td>
<td>6465</td>
<td>27.20</td>
<td>175835</td>
<td>76.01</td>
<td>66.20</td>
<td>70.79</td>
<td>1.90</td>
</tr>
<tr>
<td>5</td>
<td>9.80</td>
<td>6465</td>
<td>27.20</td>
<td>175835</td>
<td>76.32</td>
<td>66.20</td>
<td>70.79</td>
<td>1.94</td>
</tr>
<tr>
<td>6</td>
<td>9.70</td>
<td>6465</td>
<td>27.20</td>
<td>175835</td>
<td>76.63</td>
<td>66.20</td>
<td>70.92</td>
<td>2.06</td>
</tr>
<tr>
<td>7</td>
<td>9.77</td>
<td>6465</td>
<td>27.20</td>
<td>175835</td>
<td>76.01</td>
<td>66.20</td>
<td>70.86</td>
<td>1.90</td>
</tr>
<tr>
<td>8</td>
<td>15.61</td>
<td>15693</td>
<td>34.35</td>
<td>539069</td>
<td>88.13</td>
<td>71.70</td>
<td>75.57</td>
<td>1.50</td>
</tr>
<tr>
<td>9</td>
<td>15.11</td>
<td>18834</td>
<td>28.10</td>
<td>529247</td>
<td>83.68</td>
<td>70.40</td>
<td>72.55</td>
<td>2.26</td>
</tr>
<tr>
<td>10</td>
<td>13.46</td>
<td>14982</td>
<td>32.70</td>
<td>489919</td>
<td>95.82</td>
<td>73.10</td>
<td>76.45</td>
<td>2.90</td>
</tr>
<tr>
<td>11</td>
<td>4.47</td>
<td>1543</td>
<td>31.77</td>
<td>49026</td>
<td>59.48</td>
<td>53.20</td>
<td>61.14</td>
<td>0.41</td>
</tr>
<tr>
<td>12</td>
<td>6.14</td>
<td>4597</td>
<td>53.41</td>
<td>245503</td>
<td>79.50</td>
<td>68.60</td>
<td>73.70</td>
<td>1.09</td>
</tr>
<tr>
<td>13</td>
<td>15.08</td>
<td>14367</td>
<td>32.50</td>
<td>466936</td>
<td>91.90</td>
<td>72.80</td>
<td>77.18</td>
<td>4.10</td>
</tr>
<tr>
<td>14</td>
<td>15.08</td>
<td>14367</td>
<td>32.50</td>
<td>466936</td>
<td>91.90</td>
<td>72.80</td>
<td>77.18</td>
<td>4.10</td>
</tr>
<tr>
<td>15</td>
<td>15.39</td>
<td>14867</td>
<td>32.50</td>
<td>483182</td>
<td>91.90</td>
<td>72.80</td>
<td>77.51</td>
<td>4.13</td>
</tr>
<tr>
<td>16</td>
<td>14.93</td>
<td>14100</td>
<td>32.50</td>
<td>458240</td>
<td>91.90</td>
<td>72.80</td>
<td>76.93</td>
<td>4.07</td>
</tr>
<tr>
<td>17</td>
<td>14.61</td>
<td>13573</td>
<td>32.50</td>
<td>441127</td>
<td>91.90</td>
<td>72.80</td>
<td>76.80</td>
<td>4.02</td>
</tr>
<tr>
<td>18</td>
<td>15.24</td>
<td>14575</td>
<td>32.50</td>
<td>473695</td>
<td>91.90</td>
<td>72.80</td>
<td>77.29</td>
<td>4.12</td>
</tr>
<tr>
<td>19</td>
<td>15.24</td>
<td>14575</td>
<td>32.50</td>
<td>473695</td>
<td>91.90</td>
<td>72.80</td>
<td>77.29</td>
<td>4.12</td>
</tr>
<tr>
<td>20</td>
<td>15.39</td>
<td>14809</td>
<td>32.50</td>
<td>481293</td>
<td>91.90</td>
<td>72.80</td>
<td>77.18</td>
<td>4.12</td>
</tr>
<tr>
<td>21</td>
<td>14.45</td>
<td>13325</td>
<td>32.50</td>
<td>433061</td>
<td>91.90</td>
<td>72.80</td>
<td>76.78</td>
<td>3.98</td>
</tr>
<tr>
<td>22</td>
<td>14.77</td>
<td>13865</td>
<td>32.50</td>
<td>450611</td>
<td>91.90</td>
<td>72.80</td>
<td>76.88</td>
<td>4.04</td>
</tr>
<tr>
<td>23</td>
<td>14.77</td>
<td>13865</td>
<td>32.50</td>
<td>450611</td>
<td>91.90</td>
<td>72.80</td>
<td>76.88</td>
<td>4.04</td>
</tr>
<tr>
<td>24</td>
<td>15.08</td>
<td>14367</td>
<td>32.50</td>
<td>466936</td>
<td>91.90</td>
<td>72.80</td>
<td>77.18</td>
<td>4.10</td>
</tr>
<tr>
<td>25</td>
<td>14.45</td>
<td>13387</td>
<td>32.50</td>
<td>435077</td>
<td>91.90</td>
<td>72.80</td>
<td>76.80</td>
<td>4.00</td>
</tr>
<tr>
<td>26</td>
<td>11.13</td>
<td>11848</td>
<td>40.10</td>
<td>475118</td>
<td>95.37</td>
<td>70.00</td>
<td>73.74</td>
<td>1.92</td>
</tr>
<tr>
<td>27</td>
<td>12.12</td>
<td>20213</td>
<td>40.10</td>
<td>810555</td>
<td>94.15</td>
<td>70.00</td>
<td>74.82</td>
<td>2.40</td>
</tr>
<tr>
<td>28</td>
<td>12.00</td>
<td>19070</td>
<td>40.10</td>
<td>764707</td>
<td>95.25</td>
<td>70.00</td>
<td>74.68</td>
<td>2.40</td>
</tr>
<tr>
<td>29</td>
<td>11.05</td>
<td>11680</td>
<td>40.10</td>
<td>468368</td>
<td>95.82</td>
<td>70.00</td>
<td>73.54</td>
<td>1.99</td>
</tr>
<tr>
<td>30</td>
<td>12.12</td>
<td>20213</td>
<td>40.10</td>
<td>810555</td>
<td>94.15</td>
<td>70.00</td>
<td>74.82</td>
<td>2.40</td>
</tr>
<tr>
<td>31</td>
<td>10.71</td>
<td>10095</td>
<td>40.10</td>
<td>404810</td>
<td>93.90</td>
<td>70.00</td>
<td>72.86</td>
<td>1.83</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>12.06</td>
<td>19724</td>
<td>40.10</td>
<td>790932</td>
<td>94.66</td>
<td>70.00</td>
<td>74.75</td>
<td>2.40</td>
</tr>
<tr>
<td>33</td>
<td>12.29</td>
<td>22466</td>
<td>40.10</td>
<td>900901</td>
<td>94.81</td>
<td>70.00</td>
<td>75.14</td>
<td>2.50</td>
</tr>
<tr>
<td>34</td>
<td>12.29</td>
<td>22466</td>
<td>40.10</td>
<td>900901</td>
<td>94.81</td>
<td>70.00</td>
<td>75.14</td>
<td>2.50</td>
</tr>
<tr>
<td>35</td>
<td>12.29</td>
<td>22466</td>
<td>40.10</td>
<td>900901</td>
<td>94.81</td>
<td>70.00</td>
<td>75.14</td>
<td>2.50</td>
</tr>
<tr>
<td>36</td>
<td>11.89</td>
<td>17235</td>
<td>40.10</td>
<td>691124</td>
<td>96.05</td>
<td>70.00</td>
<td>74.59</td>
<td>2.40</td>
</tr>
<tr>
<td>37</td>
<td>11.89</td>
<td>17235</td>
<td>40.10</td>
<td>691124</td>
<td>96.05</td>
<td>70.00</td>
<td>74.59</td>
<td>2.40</td>
</tr>
<tr>
<td>38</td>
<td>10.85</td>
<td>10964</td>
<td>40.10</td>
<td>439646</td>
<td>95.37</td>
<td>70.00</td>
<td>73.12</td>
<td>1.92</td>
</tr>
<tr>
<td>39</td>
<td>11.30</td>
<td>13791</td>
<td>40.10</td>
<td>553036</td>
<td>95.57</td>
<td>70.00</td>
<td>73.84</td>
<td>2.09</td>
</tr>
<tr>
<td>40</td>
<td>12.46</td>
<td>28527</td>
<td>40.80</td>
<td>1163888</td>
<td>96.43</td>
<td>70.00</td>
<td>75.93</td>
<td>2.63</td>
</tr>
<tr>
<td>41</td>
<td>11.52</td>
<td>14408</td>
<td>40.10</td>
<td>577741</td>
<td>95.89</td>
<td>70.00</td>
<td>74.25</td>
<td>1.99</td>
</tr>
<tr>
<td>42</td>
<td>11.64</td>
<td>15444</td>
<td>40.10</td>
<td>602658</td>
<td>95.78</td>
<td>70.00</td>
<td>74.37</td>
<td>1.99</td>
</tr>
<tr>
<td>43</td>
<td>11.58</td>
<td>15029</td>
<td>40.10</td>
<td>602658</td>
<td>95.78</td>
<td>70.00</td>
<td>74.31</td>
<td>1.99</td>
</tr>
<tr>
<td>44</td>
<td>11.58</td>
<td>15029</td>
<td>40.10</td>
<td>602658</td>
<td>95.78</td>
<td>70.00</td>
<td>74.31</td>
<td>1.99</td>
</tr>
<tr>
<td>45</td>
<td>11.64</td>
<td>15029</td>
<td>40.10</td>
<td>602658</td>
<td>95.57</td>
<td>70.00</td>
<td>74.36</td>
<td>1.99</td>
</tr>
<tr>
<td>46</td>
<td>12.4</td>
<td>22998</td>
<td>40.10</td>
<td>922200</td>
<td>93.08</td>
<td>70.00</td>
<td>75.31</td>
<td>2.47</td>
</tr>
<tr>
<td>47</td>
<td>5.73</td>
<td>13885</td>
<td>31.60</td>
<td>438777</td>
<td>98.30</td>
<td>65.00</td>
<td>72.31</td>
<td>1.15</td>
</tr>
<tr>
<td>49</td>
<td>5.37</td>
<td>11357</td>
<td>31.60</td>
<td>358882</td>
<td>91.04</td>
<td>65.00</td>
<td>70.96</td>
<td>1.02</td>
</tr>
<tr>
<td>50</td>
<td>5.37</td>
<td>11357</td>
<td>31.60</td>
<td>358882</td>
<td>91.04</td>
<td>65.00</td>
<td>70.96</td>
<td>1.02</td>
</tr>
<tr>
<td>51</td>
<td>12.32</td>
<td>10831</td>
<td>26.60</td>
<td>288103</td>
<td>90.84</td>
<td>68.00</td>
<td>71.78</td>
<td>2.77</td>
</tr>
<tr>
<td>52</td>
<td>14.90</td>
<td>19805</td>
<td>29.00</td>
<td>574344</td>
<td>89.29</td>
<td>70.40</td>
<td>75.00</td>
<td>2.99</td>
</tr>
<tr>
<td>53</td>
<td>5.90</td>
<td>1748</td>
<td>41.50</td>
<td>72549</td>
<td>60.90</td>
<td>62.30</td>
<td>69.01</td>
<td>1.54</td>
</tr>
<tr>
<td>54</td>
<td>5.49</td>
<td>1291</td>
<td>41.50</td>
<td>53591</td>
<td>56.40</td>
<td>68.87</td>
<td>69.01</td>
<td>1.54</td>
</tr>
<tr>
<td>55</td>
<td>4.03</td>
<td>6385</td>
<td>50.30</td>
<td>321171</td>
<td>57.53</td>
<td>65.00</td>
<td>70.59</td>
<td>1.13</td>
</tr>
<tr>
<td>56</td>
<td>4.19</td>
<td>6914</td>
<td>50.30</td>
<td>347759</td>
<td>60.55</td>
<td>65.00</td>
<td>70.78</td>
<td>1.20</td>
</tr>
<tr>
<td>57</td>
<td>12.65</td>
<td>15945</td>
<td>31.50</td>
<td>502262</td>
<td>95.46</td>
<td>72.00</td>
<td>76.96</td>
<td>2.55</td>
</tr>
<tr>
<td>58</td>
<td>12.38</td>
<td>13982</td>
<td>31.50</td>
<td>440431</td>
<td>96.32</td>
<td>72.00</td>
<td>76.77</td>
<td>2.44</td>
</tr>
<tr>
<td>59</td>
<td>12.75</td>
<td>16734</td>
<td>32.10</td>
<td>537151</td>
<td>96.3</td>
<td>72.00</td>
<td>77.06</td>
<td>2.55</td>
</tr>
<tr>
<td>60</td>
<td>15.76</td>
<td>17520</td>
<td>36.10</td>
<td>632486</td>
<td>92.37</td>
<td>71.70</td>
<td>76.29</td>
<td>1.55</td>
</tr>
<tr>
<td>61</td>
<td>14.22</td>
<td>20006</td>
<td>34.60</td>
<td>692209</td>
<td>80.48</td>
<td>72.50</td>
<td>77.17</td>
<td>1.59</td>
</tr>
<tr>
<td>62</td>
<td>14.22</td>
<td>20006</td>
<td>34.60</td>
<td>692209</td>
<td>80.48</td>
<td>72.50</td>
<td>77.17</td>
<td>1.58</td>
</tr>
<tr>
<td>63</td>
<td>11.17</td>
<td>13755</td>
<td>43.90</td>
<td>603855</td>
<td>89.13</td>
<td>69.20</td>
<td>75.47</td>
<td>1.95</td>
</tr>
<tr>
<td>64</td>
<td>11.23</td>
<td>16716</td>
<td>33.70</td>
<td>563335</td>
<td>90.66</td>
<td>73.20</td>
<td>77.29</td>
<td>2.30</td>
</tr>
<tr>
<td>65</td>
<td>11.06</td>
<td>16278</td>
<td>33.70</td>
<td>548559</td>
<td>90.37</td>
<td>73.20</td>
<td>76.90</td>
<td>2.30</td>
</tr>
<tr>
<td>66</td>
<td>11.54</td>
<td>18838</td>
<td>35.20</td>
<td>663111</td>
<td>93.03</td>
<td>73.20</td>
<td>77.85</td>
<td>2.40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>67</td>
<td>15.83</td>
<td>18962</td>
<td>27.30</td>
<td>517649</td>
<td>87.23</td>
<td>72.70</td>
<td>77.17</td>
<td>5.20</td>
</tr>
<tr>
<td>68</td>
<td>15.86</td>
<td>18953</td>
<td>29.30</td>
<td>555335</td>
<td>87.13</td>
<td>72.70</td>
<td>77.17</td>
<td>5.20</td>
</tr>
<tr>
<td>69</td>
<td>10.64</td>
<td>16149</td>
<td>31.50</td>
<td>508697</td>
<td>91.12</td>
<td>72.00</td>
<td>76.59</td>
<td>2.09</td>
</tr>
<tr>
<td>70</td>
<td>4.81</td>
<td>5284</td>
<td>46.20</td>
<td>244100</td>
<td>37.4</td>
<td>60.2</td>
<td>69.99</td>
<td>0.23</td>
</tr>
</tbody>
</table>

Sources: Data for all of the variables except Disability-Adjusted Life Expectancy (DALE) were obtained from SIMA (World Bank’s internal database system). The DALE estimate for the population of each country was found in the World Health Report, Annex Table 5, of the World Health Organisation.
### Appendix 6.E: Regression Results for the Full Sample Using OLS and VWLS

#### reg regco poll over65 gnp gini educ dale

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>Number of obs = 70</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>3.99686254</td>
<td>6</td>
<td>.666143757</td>
<td>F( 6, 63) = 1.16</td>
</tr>
<tr>
<td>Residual</td>
<td>36.1019116</td>
<td>63</td>
<td>.573046215</td>
<td>Prob &gt; F = 0.3377</td>
</tr>
<tr>
<td>Total</td>
<td>40.0987741</td>
<td>69</td>
<td>.581141654</td>
<td>R-squared = 0.0997</td>
</tr>
</tbody>
</table>

| regco | Coef.     | Std. Err. | t     | P>|t|  | 95% Conf. Interval |
|-------|-----------|-----------|-------|------|------------------|
| poll  | -.0031723 | .003138   | -1.011| 0.316| -.0094432 .0030986|
| over65| .0117826  | .0617333  | 0.191 | 0.849| -.1115815 .1351468|
| gnp   | .0000198  | .0000273  | 0.724 | 0.472| -.0000349 .0000744|
| gini  | .0294887  | .0185537  | 1.589 | 0.117| -.0075879 .0665652|
| educ  | -.011089  | .0122062  | -0.908| 0.367| -.0354811 .0133032|
| dale  | .0030789  | .0522493  | 0.059 | 0.953| -.1013329 .1074907|
| _cons | .1696556  | 2.989451  | 0.057 | 0.955| -5.804285 6.143596|

#### reg regco poll over65 gnp dist educ dale

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>Number of obs = 70</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>3.64420902</td>
<td>6</td>
<td>.60736817</td>
<td>F( 6, 63) = 1.05</td>
</tr>
<tr>
<td>Residual</td>
<td>36.4545651</td>
<td>63</td>
<td>.57864389</td>
<td>Prob &gt; F = 0.4023</td>
</tr>
<tr>
<td>Total</td>
<td>40.0987741</td>
<td>69</td>
<td>.581141654</td>
<td>R-squared = 0.0909</td>
</tr>
</tbody>
</table>

| regco | Coef.     | Std. Err. | t     | P>|t|  | 95% Conf. Interval |
|-------|-----------|-----------|-------|------|------------------|
| poll  | -.0028316 | .003138   | -0.903| 0.370| -.0090103 .003437|
| over65| -.0060344 | .0581366  | -0.104| 0.918| -.1222111 .1101424|
| gnp   | -.000055  | .000071   | -0.775| 0.441| -.0000197 .0000869|
| dist  | 2.10e-06 | 1.53e-06  | 1.376 | 0.174| -9.51e-07 5.15e-06|
| educ  | -.01555  | .0120439  | -1.291| 0.201| -.0396177 .0085178|
| dale  | .0276085 | .0494173  | 0.559 | 0.578| -.071144 1.263611|
| _cons | .090421  | 3.002442  | 0.030 | 0.976| -5.90948 6.090322|

303
. *vwls regco poll over65 gnp gini educ dale, sd(se)*

Variance-weighted least-squares regression

|       | Coef.     | Std. Err. | z   | P>|z|  | [95% Conf. Interval] |
|-------|-----------|-----------|-----|------|---------------------|
| poll  | 0.001852  | 0.0006687 | 2.770 | 0.006 | 0.0005414 - 0.0031626 |
| over65| -0.0838304| 0.0167197 | -5.014 | 0.000 | -0.1166004 - 0.0510604 |
| gnp   | 0.000402  | 8.65e-06  | 4.648 | 0.000 | 0.0000233 - 0.0000572 |
| gini  | -0.0042486| 0.0035616 | -1.193 | 0.233 | -0.0112292 - 0.002732 |
| educ  | -0.0019518| 0.0040647 | -0.480 | 0.631 | -0.0099185 - 0.0060149 |
| dale  | 0.0341614 | 0.0122922 | 2.779 | 0.005 | 0.0100691 - 0.0582537 |
| _cons | -1.365648 | 0.6996823 | -1.952 | 0.051 | -2.737 - 0.005704 |

Number of obs = 70
Model chi2(6) = 47.57
Prob > chi2 = 0.0000

. *vwls regco poll over65 gnp dist educ dale, sd(se)*

Variance-weighted least-squares regression

|       | Coef.     | Std. Err. | z   | P>|z|  | [95% Conf. Interval] |
|-------|-----------|-----------|-----|------|---------------------|
| poll  | 0.0016237 | 0.0005708 | 2.421 | 0.015 | 0.0003091 - 0.0029384 |
| over65| -0.0845519| 0.0154389 | -5.477 | 0.000 | -0.1148116 - 0.0542921 |
| gnp   | -0.000605 | 0.000155  | 3.912 | 0.000 | 0.0000302 - 0.0000908 |
| dist  | -5.59e-07 | 3.12e-07  | -1.791 | 0.073 | -1.17e-06 - 5.27e-08  |
| educ  | -0.0014427| 0.0040749 | -0.354 | 0.723 | -0.0094294 - 0.0065441 |
| dale  | 0.0300672 | 0.0117339 | 2.562 | 0.010 | 0.0070692 - 0.0530653 |
| _cons | -1.257962 | 0.6975314 | -1.803 | 0.071 | -2.625098 - 1.091746 |
Appendix 6.F: Regression Results for a Random Sample of Sample Combinations.

In the regression results supplied below, each country enters with only one observation.

```
. ml model d0 dprog1 (regco = poll over65 gnp gini educ dale ) / tau2 if c1>0
Number of obs    =     21
Wald chi2(6)    =   6.08
Log likelihood = -41.361076  Prob > chi2    =   0.4147

| Coef.  | Std. Err.  | z     | P>|z|       | [95% Conf. Interval] |
|--------|------------|-------|----------|----------------------|
| poll   | -.0008634  | .0016022 | -0.539  | 0.590 | -.0040036 .0022769 |
| over65 | -.0086392  | .0356966 | -0.242  | 0.809 | -.0786033 .0613248 |
| gnp    | .0000103   | .000206  | 0.498   | 0.618 | -.0000301 .0000507 |
| gini   | .0129299   | .0102547 | 1.261   | 0.207 | -.0071689 .0330287 |
| educ   | -.0038472  | .0064883 | -0.593  | 0.553 | -.016564  .0088696 |
| dale   | .0074275   | .027462  | 0.270   | 0.787 | -.046397  .061252  |
| _cons  | -.0765322  | 1.354843 | -0.056  | 0.955 | -2.731975 2.578911 |

Log likelihood = -59.636477  Prob > chi2    =   0.0658
```

```
. ml model d0 dprog2 (regco = poll over65 gnp gini educ dale ) / tau2 if c2>0
Number of obs    =     21
Wald chi2(6)    =  11.83
Log likelihood = -59.636477  Prob > chi2    =   0.0058

| Coef.  | Std. Err.  | z     | P>|z|       | [95% Conf. Interval] |
|--------|------------|-------|----------|----------------------|
| poll   | -.0005737  | .0026475 | -0.217  | 0.828 | -.0057627 .0046153 |
| over65 | -.0957259  | .053393  | -1.793  | 0.073 | -.2003742 .0089224 |
| gnp    | .0000858   | .000347  | 2.472   | 0.013 | .000178  .0001538 |
| gini   | -.0028934  | .0155448 | -0.186  | 0.852 | -.0333607 .0275739 |
| educ   | -.0228124  | .0108283 | -2.107  | 0.035 | -.0440355 .0015892 |
| dale   | .030321    | .0346143 | 0.876   | 0.381 | -.0375217 .0981637 |
| _cons  | .518916    | 1.860709 | 0.279   | 0.780 | -3.128007 4.165839 |
```
. ml model d0 d0prog3 (regco = poll over65 gnp gini educ dale) / tau2 if c3>0
Number of obs = 21
Wald chi2(6) = 12.81
Log likelihood = -46.552821 Prob > chi2 = 0.0462

| regco | Coef.  | Std. Err. | z   | P>|z|  | [95% Conf. Interval] |
|-------|--------|-----------|-----|------|----------------------|
| poll  | -0.0004785 | 0.0013978 | -0.342 | 0.732 | -0.0032181 - 0.0022611 |
| over65| -0.0058429 | 0.0380754 | -0.153 | 0.878 | -0.0804693 - 0.0687835 |
| gnp   | 8.93e-06    | 0.0000215 | 0.416 | 0.678 | -0.0000332 - 0.000051 |
| gini  | 0.0117473   | 0.0099307 | 1.183 | 0.237 | -0.0077166 - 0.0311212 |
| educ  | -0.0164378  | 0.0070205 | -2.341 | 0.019 | -0.0301977 - 0.026778 |
| dale  | 0.0274064   | 0.027457  | 0.998 | 0.318 | -0.0264083 - 0.0812211 |
| _cons | -0.5030326  | 1.31468   | -0.383 | 0.702 | -3.079759 - 2.073694 |

tau2
| _cons | 0.0669139   | 0.361628  | 1.850 | 0.064 | -0.039638 - 0.137916 |

. ml model d0 d0prog4 (regco = poll over65 gnp gini educ dale) / tau2 if c4>0
Number of obs = 21
Wald chi2(6) = 12.82
Log likelihood = -56.603155 Prob > chi2 = 0.0460

| regco | Coef.  | Std. Err. | z   | P>|z|  | [95% Conf. Interval] |
|-------|--------|-----------|-----|------|----------------------|
| poll  | 0.0026665  | 0.002568 | 1.038 | 0.299 | -0.0023666 - 0.0076996 |
| over65| 0.0534209  | 0.047811 | 1.117 | 0.264 | -0.0402869 - 0.1471288 |
| gnp   | 0.000069    | 0.000031 | 2.222 | 0.026 | 8.15e-06 - 0.001298 |
| gini  | 0.022429    | 0.014020 | 1.600 | 0.110 | -0.0050498 - 0.0499079 |
| educ  | -0.0131625  | 0.0086045 | -1.530 | 0.126 | -0.0300271 - 0.037021 |
| dale  | -0.0188534  | 0.0313242 | -0.602 | 0.547 | -0.0802477 - 0.0425409 |
| _cons | 0.6598235   | 1.636398 | 0.403 | 0.687 | -2.547458 - 3.867105 |

tau2
| _cons | 0.2100956  | 0.0834808 | 2.517 | 0.012 | 0.464763 - 0.373715 |

. ml model d0 d0prog5 (regco = poll over65 gnp gini educ dale) / tau2 if c5>0
Number of obs = 21
Wald chi2(6) = 20.96
Log likelihood = -39.845363 Prob > chi2 = 0.0019

| regco | Coef.  | Std. Err. | z   | P>|z|  | [95% Conf. Interval] |
|-------|--------|-----------|-----|------|----------------------|
| poll  | 0.0006733  | 0.0014577 | 0.462 | 0.644 | -0.0021837 - 0.0035302 |
| over65| -0.1097985 | 0.0293858 | -3.736 | 0.000 | -1.673936 - 0.0522033 |
| gnp   | 0.0000457  | 0.0001888 | 2.433 | 0.015 | 8.89e-06 - 0.000826 |
| gini  | -0.005942  | 0.0081729 | -0.727 | 0.467 | -0.0219607 - 0.0100767 |
| educ  | -0.0086456 | 0.0061046 | -1.416 | 0.157 | -0.0206103 - 0.0033191 |
| dale  | 0.0405547  | 0.0228899 | 1.772 | 0.076 | -0.0043087 - 0.0854181 |
| _cons | -0.8398555 | 1.200993  | -0.699 | 0.484 | -3.193758 - 1.514047 |

tau2
| _cons | 0.0551182  | 0.0249164 | 2.212 | 0.027 | 0.0062829 - 0.1039535 |

306
```
.ml model d0 d0prog0 (regco = poll over65 gnp gini educ dale) / tau2 if c6>0
Number of obs = 21
Wald chi2(6) = 10.36
Prob > chi2 = 0.1105
Log likelihood = -60.052485

               | Coef.     | Std. Err. |      z    |   P>|z|     |      [95% Conf. Interval]
-------------|-----------|-----------|-----------|--------|--------------------------
poll         | -0.0011176| 0.002853  | -0.392   | 0.695  | -0.0067094 .0044742
over65       | 0.0157608 | 0.0593338 | 0.266    | 0.791  | -1.005314 .1320529
gnp          | 0.0000434 | 0.0000409 | 1.062    | 0.288  | -0.0000367 .0001236
egini        | 0.020833  | 0.0160285 | 1.300    | 0.194  | -0.0105823 .0522484
educ         | -0.0217648| 0.0107806 | -2.019   | 0.043  | -0.0428944 .006352
dale         | 0.0076838 | 0.0362123 | 0.212    | 0.832  | -0.063291 .0786586
_cons        | 0.5178108 | 1.89681   | 0.273    | 0.785  | -3.199869 4.235491

.ml model d0 d0prog7 (regco = poll over65 gnp gini educ dale) / tau2 if c7>0
Number of obs = 21
Wald chi2(6) = 11.10
Log likelihood = -50.665526
Prob > chi2 = 0.0852

               | Coef.     | Std. Err. |      z    |   P>|z|     |      [95% Conf. Interval]
poll         | -0.0008796| 0.0017237 | -0.510   | 0.610  | -0.0042581 .0024988
over65       | -0.020491 | 0.0450024 | -0.455   | 0.649  | -0.1086942 .0677121
gnp          | 0.00004   | 0.0000242 | 1.653    | 0.098  | -7.43e-06 .0000874
egini        | 0.021933  | 0.0125281 | 1.300    | 0.194  | -0.0033613 .0457478
educ         | -0.0060512| 0.0082835 | -0.731   | 0.465  | -0.0748827 .0602445
dale         | -0.0073191| 0.0344719 | -0.212   | 0.832  | -0.0748827 .0602445
_cons        | 0.6446868 | 1.638074  | 0.394    | 0.694  | -2.565879 3.855253

.ml model d0 d0prog8 (regco = poll over65 gnp gini educ dale) / tau2 if c8>0
Number of obs = 21
Wald chi2(6) = 11.56
Log likelihood = -57.789363
Prob > chi2 = 0.0724

               | Coef.     | Std. Err. |      z    |   P>|z|     |      [95% Conf. Interval]
poll         | 0.0021584 | 0.002653  | 0.814    | 0.416  | -0.0030414 .0073582
over65       | -0.0557361| 0.0490866 | -1.135   | 0.256  | -1.519441 .0404719
gnp          | 0.000095  | 0.000329  | 2.886    | 0.004  | 0.000305 0.001596
egini        | 0.0022487 | 0.0143583 | 0.157    | 0.876  | -0.025893 0.030903
educ         | -0.0144162| 0.009787  | -1.445   | 0.149  | -0.033974 .0051417
dale         | 0.0154026 | 0.03134   | 0.491    | 0.623  | -0.0460227 .0768278
_cons        | -0.0106618| 1.692653  | -0.006   | 0.995  | -3.3282 3.306876
```

. ml model d0 d0prog9 (regco = poll over65 gnp gini educ dale) / tau2 if c9>0
Number of obs = 21
Wald chi2(6) = 20.60
Log likelihood = -36.620218
Prob > chi2 = 0.0022

| regco  | Coef.  | Std. Err. | z     | P>|z|  | [95% Conf. Interval] |
|--------|--------|-----------|-------|------|----------------------|
| poll   | 0.000538 | 0.001138 | 0.473 | 0.636 | -0.0016928 to 0.0027694 |
| over65 | -0.0262215 | 0.0278982 | -0.940 | 0.347 | -0.0809011 to 0.028458 |
| gnp    | 0.000017 | 0.000148 | 1.153 | 0.249 | -0.0000119 to 0.000046 |
| gini   | 0.0044916 | 0.0068194 | 0.659 | 0.510 | -0.0088743 to 0.017854 |
| educ   | -0.0163131 | 0.0052966 | -3.080 | 0.002 | -0.0266943 to -0.0059319 |
| dale   | 0.0362147 | 0.0190977 | 1.896 | 0.058 | -0.0012162 to 0.0763655 |
| _cons  | -0.8936275 | 0.9963862 | -0.897 | 0.370 | -2.846509 to 1.059254 |

| tau2   | Coef.  | Std. Err. | z     | P>|z|  | [95% Conf. Interval] |
|--------|--------|-----------|-------|------|----------------------|
| _cons  | 0.0270157 | 0.0143073 | 1.888 | 0.059 | -0.0010261 to 0.0550575 |

. ml model d0 d0prog10 (regco = poll over65 gnp gini educ dale) / tau2 if c10>0
Number of obs = 21
Wald chi2(6) = 11.38
Log likelihood = -59.331796
Prob > chi2 = 0.0772

| regco  | Coef.  | Std. Err. | z     | P>|z|  | [95% Conf. Interval] |
|--------|--------|-----------|-------|------|----------------------|
| poll   | -0.0004563 | 0.0025717 | -0.177 | 0.859 | -0.0054968 to 0.0045842 |
| over65 | 0.0150591 | 0.0538812 | 0.279 | 0.780 | -0.0905461 to 0.1206643 |
| gnp    | 0.0000571 | 0.0000308 | 1.858 | 0.063 | -3.13e-06 to 0.0001174 |
| gini   | 0.0164117 | 0.0159063 | 1.032 | 0.302 | -0.0476411 to 0.0475875 |
| educ   | 0.0020872 | 0.0102959 | -1.015 | 0.309 | -0.0261767 to 0.036363 |
| dale   | -0.0036402 | 0.0098769 | -0.105 | 0.916 | -0.0227329 to 0.0154525 |
| _cons  | 1.174953 | 1.795952 | 0.654 | 0.513 | -2.345049 to 4.694956 |

| tau2   | Coef.  | Std. Err. | z     | P>|z|  | [95% Conf. Interval] |
|--------|--------|-----------|-------|------|----------------------|
| _cons  | 0.264388 | 0.1024882 | 2.580 | 0.010 | 0.0635148 to 0.4652612 |

. ml model d0 d0prog11 (regco = poll over65 gnp gini educ dale) / tau2 if c11>0
Number of obs = 21
Wald chi2(6) = 14.21
Log likelihood = -46.567254
Prob > chi2 = 0.0273

| regco  | Coef.  | Std. Err. | z     | P>|z|  | [95% Conf. Interval] |
|--------|--------|-----------|-------|------|----------------------|
| poll   | -0.0005675 | 0.001445 | -0.393 | 0.695 | -0.0033995 to 0.0022646 |
| over65 | -0.0946786 | 0.034787 | -2.722 | 0.007 | -0.16286 to -0.0264973 |
| gnp    | 0.000355 | 0.000225 | 1.574 | 0.116 | -8.71e-06 to 0.0000797 |
| gini   | 0.008488 | 0.0098769 | 0.869 | 0.383 | -0.0185097 to 0.0202702 |
| educ   | -0.006141 | 0.0071078 | -1.353 | 0.176 | -0.0235452 to 0.004317 |
| dale   | 0.038468 | 0.0273919 | 1.404 | 0.160 | -0.0152191 to 0.0692155 |
| _cons  | -0.7223563 | 1.318953 | -0.548 | 0.584 | -3.307457 to 1.862745 |

| tau2   | Coef.  | Std. Err. | z     | P>|z|  | [95% Conf. Interval] |
|--------|--------|-----------|-------|------|----------------------|
| _cons  | 0.0714972 | 0.0340504 | 2.100 | 0.036 | 0.0047597 to 1.382347 |

308
. `ml model d0 d0progl2 (regco = poll over65 gnp gini educ dale) / tau2 if cl2 > 0`

Number of obs = 21
Wald chi2(6) = 13.37
Prob > chi2 = 0.0376

Log likelihood = -60.762544

| regco | Coef.  | Std. Err. | z     | P>|z|   | [95% Conf. Interval] |
|-------|--------|-----------|-------|------|-----------------------|
| poll  | 0.0023236 | 0.002649  | 0.877 | 0.380 | -.0028682 to 0.0075155 |
| over65| 0.0138613  | 0.0516725 | 0.268 | 0.789 | -.0874149 to 0.1151375 |
| gnp   | 0.0000739  | 0.000036  | 2.056 | 0.040 | 3.45e-06 to 0.001444  |
| gini  | 0.0103033  | 0.0142352 | 0.724 | 0.469 | -.0175972 to 0.0382038 |
| educ  | -0.023356  | 0.0098721 | -2.366 | 0.018 | -.0427051 to -.004007  |
| dale  | 0.0107396  | 0.0310045 | 0.346 | 0.729 | -.0500281 to 0.0715073 |
| _cons | 0.1546618  | 1.654462  | 0.093 | 0.926 | -3.088025 to 3.397349  |

. `ml model d0 d0progl3 (regco = poll over65 gnp gini educ dale) / tau2 if cl3 > 0`

Number of obs = 21
Wald chi2(6) = 6.75
Prob > chi2 = 0.3445

Log likelihood = -53.687934

| regco | Coef.  | Std. Err. | z     | P>|z|   | [95% Conf. Interval] |
|-------|--------|-----------|-------|------|-----------------------|
| poll  | -0.0008195 | 0.0018503 | -0.443 | 0.658 | -.004446 to 0.002807  |
| over65| -0.0352514 | 0.0395734 | -0.891 | 0.373 | -0.1128139 to 0.042311 |
| gnp   | 0.0000246  | 0.0000248 | 0.993 | 0.321 | -0.00024 to 0.000733  |
| gini  | 0.0117938  | 0.0116333 | 1.014 | 0.311 | -0.011007 to 0.0345947 |
| educ  | -0.0045706 | 0.0071821 | -0.636 | 0.525 | -0.0186472 to 0.009506 |
| dale  | 0.0036871  | 0.0319272 | 0.115 | 0.908 | -.058889 to 0.0662632 |
| _cons | 0.3479319  | 1.575462  | 0.221 | 0.825 | -2.739917 to 3.435781  |

. `ml model d0 d0progl4 (regco = poll over65 gnp gini educ dale) / tau2 if cl4 > 0`

Number of obs = 21
Wald chi2(6) = 10.32
Prob > chi2 = 0.1119

Log likelihood = -58.404488

| regco | Coef.  | Std. Err. | z     | P>|z|   | [95% Conf. Interval] |
|-------|--------|-----------|-------|------|-----------------------|
| poll  | 0.0002928 | 0.0027661 | 0.106 | 0.916 | -.0051286 to 0.0057142 |
| over65| -0.0668596 | 0.0534701 | -1.250 | 0.211 | -.1716591 to 0.03794  |
| gnp   | 0.00087   | 0.000039  | 2.228 | 0.026 | 0.0000105 to 0.0001635 |
| gini  | 0.0055993  | 0.0147901 | 0.379 | 0.705 | -0.0233886 to 0.0345873 |
| educ  | -0.018301  | 0.0100758 | -1.816 | 0.069 | -0.0380492 to 0.0014471 |
| dale  | 0.0208245  | 0.033586  | 0.624 | 0.532 | -0.0445571 to 0.0862061 |
| _cons | 0.2107019  | 1.785615  | 0.118 | 0.906 | -3.289038 to 3.710442  |

. `ml model d0 d0progl4 (regco = poll over65 gnp gini educ dale) / tau2 if cl4 > 0`

Number of obs = 21
Wald chi2(6) = 10.32
Prob > chi2 = 0.1119

Log likelihood = -58.404488

| regco | Coef.  | Std. Err. | z     | P>|z|   | [95% Conf. Interval] |
|-------|--------|-----------|-------|------|-----------------------|
| poll  | 0.0002928 | 0.0027661 | 0.106 | 0.916 | -.0051286 to 0.0057142 |
| over65| -0.0668596 | 0.0534701 | -1.250 | 0.211 | -.1716591 to 0.03794  |
| gnp   | 0.00087   | 0.000039  | 2.228 | 0.026 | 0.0000105 to 0.0001635 |
| gini  | 0.0055993  | 0.0147901 | 0.379 | 0.705 | -0.0233886 to 0.0345873 |
| educ  | -0.018301  | 0.0100758 | -1.816 | 0.069 | -0.0380492 to 0.0014471 |
| dale  | 0.0208245  | 0.033586  | 0.624 | 0.532 | -0.0445571 to 0.0862061 |
| _cons | 0.2107019  | 1.785615  | 0.118 | 0.906 | -3.289038 to 3.710442  |

. `ml model d0 d0progl4 (regco = poll over65 gnp gini educ dale) / tau2 if cl4 > 0`

Number of obs = 21
Wald chi2(6) = 10.32
Prob > chi2 = 0.1119

Log likelihood = -58.404488

| regco | Coef.  | Std. Err. | z     | P>|z|   | [95% Conf. Interval] |
|-------|--------|-----------|-------|------|-----------------------|
| poll  | 0.0002928 | 0.0027661 | 0.106 | 0.916 | -.0051286 to 0.0057142 |
| over65| -0.0668596 | 0.0534701 | -1.250 | 0.211 | -.1716591 to 0.03794  |
| gnp   | 0.00087   | 0.000039  | 2.228 | 0.026 | 0.0000105 to 0.0001635 |
| gini  | 0.0055993  | 0.0147901 | 0.379 | 0.705 | -0.0233886 to 0.0345873 |
| educ  | -0.018301  | 0.0100758 | -1.816 | 0.069 | -0.0380492 to 0.0014471 |
| dale  | 0.0208245  | 0.033586  | 0.624 | 0.532 | -0.0445571 to 0.0862061 |
| _cons | 0.2107019  | 1.785615  | 0.118 | 0.906 | -3.289038 to 3.710442  |
```plaintext
.ml model d0 d0progl5 (regco = poll over65 gnp gini educ dale) / tau2 if cl5>0
Number of obs = 21
Wald chi2(6) = 15.76
Prob > chi2 = 0.0151
Log likelihood = -41.859212

| regco | Coef. | Std. Err. | z    | P>|z|   | [95% Conf. Interval] |
|-------|-------|-----------|------|-------|---------------------|
| poll  | -0.0002745 | 0.0011956 | -0.230 | 0.818 | -0.0026179 -0.0020688 |
| over65| -0.0284832 | 0.0336784 | -0.846 | 0.398 | -0.0944917 0.0375253 |
| gnp   | 0.000135 | 0.000162 | 0.829 | 0.407 | -0.000184 0.0000453 |
| gini  | 0.00618 | 0.0087359 | 0.707 | 0.479 | -0.0109421 0.0233021 |
| educ  | -0.0186527 | 0.064701 | -2.883 | 0.004 | -0.0944917 0.0375253 |
| dale  | 0.0469472 | 0.0241686 | 1.682 | 0.093 | -0.0067223 0.0880167 |
| _cons | -0.8771405 | 1.15596 | -0.759 | 0.448 | -3.14278 1.388499 |

.ml model d0 d0progl6 (regco = poll over65 gnp gini educ dale) / tau2 if cl6>0
Number of obs = 21
Wald chi2(6) = 11.29
Prob > chi2 = 0.0800
Log likelihood = -58.651617

| regco | Coef. | Std. Err. | z    | P>|z|   | [95% Conf. Interval] |
|-------|-------|-----------|------|-------|---------------------|
| poll  | 0.0024846 | 0.0028603 | 0.869 | 0.385 | -0.0031216 0.0080907 |
| over65| 0.0161462 | 0.0313891 | 0.304 | 0.761 | -0.0810268 0.120395 |
| gnp   | 0.000088 | 0.000344 | 2.559 | 0.010 | -0.0000206 0.0001554 |
| gini  | 0.0216805 | 0.0159241 | 1.361 | 0.173 | -0.0095302 0.0528911 |
| educ  | -0.0135433 | 0.009639 | -1.359 | 0.174 | -0.030722 0.0059857 |
| dale  | -0.0222279 | 0.0352141 | -0.631 | 0.528 | -0.0912466 0.0467908 |
| _cons | 1.091324 | 1.841617 | 0.593 | 0.553 | -2.518179 4.700826 |

.ml model d0 d0progl7 (regco = poll over65 gnp gini educ dale) / tau2 if cl7>0
Number of obs = 21
Wald chi2(6) = 5.15
Prob > chi2 = 0.5247
Log likelihood = -45.685287

| regco | Coef. | Std. Err. | z    | P>|z|   | [95% Conf. Interval] |
|-------|-------|-----------|------|-------|---------------------|
| poll  | -0.0011067 | 0.0018063 | -0.613 | 0.540 | -0.0046471 0.0024337 |
| over65| -0.0193344 | 0.0378261 | -0.510 | 0.610 | -0.0934422 0.0548335 |
| gnp   | 0.000016 | 0.000257 | 0.622 | 0.534 | -0.0000343 0.0000663 |
| gini  | 0.0102111 | 0.110593 | 0.951 | 0.341 | -0.111548 0.032197 |
| educ  | -0.0031326 | 0.0074347 | -0.420 | 0.674 | -0.0176973 0.0114461 |
| dale  | 0.0069095 | 0.0302115 | 0.229 | 0.819 | -0.0523039 0.066123 |
| _cons | 0.0878029 | 1.481664 | 0.059 | 0.953 | -2.816204 2.99181 |
```

310
```
ml model d0 d0progl8 (regco = poll over65 gnp gini educ dale) / tau2 if cl8>0
  Number of obs = 21
  Wald chi2(6) = 13.17
  Log likelihood = -58.798355 Prob > chi2 = 0.0405

    | Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------|---------|--------------|------|--------|-----------------------------
  poll | -0.0000502   .0027792   -0.018   0.986     -.0054973    .0053969
  over65 | 0.016604     .0555866    0.299   0.765     -.0923438    .1255518
  gnp | 0.0000595   .0000397    1.497   0.134     -.0000184    .0001373
  gini | 0.148796     .0148493    1.002   0.316     -.0142245    .0443987
  educ | -0.0274776   .0102835   -2.672   0.008     -.0476328    -.0073223
  dale | 0.0131231    .0333718    0.393   0.694     -.0522845    .0785306
  _cons | 0.5365526    1.795015    0.299   0.765     -2.981612    4.054717

tau2  | _cons | 265658     .1032305    2.573   0.010     .0633298    .4679861

ml model d0 d0progl9 (regco = poll over65 gnp gini educ dale) / tau2 if c19>0
  Number of obs = 21
  Wald chi2(6) = 11.98
  Log likelihood = -48.332331 Prob > chi2 = 0.0623

    | Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------|---------|--------------|------|--------|-----------------------------
  poll | -0.0009277   .0633143   -1.466   0.143     -.1479376    .0213089
  over65 | -0.033143    .0431759   -7.679   0.000     -.1479376    .0213089
  gnp | 0.0000413   .0000253    1.632   0.103     -.0000793    .0001373
  gini | 0.112070    .012489    8.907   0.370     -.013271     .035685
  educ | -0.0107936   .0080988   -1.333   0.183     -.026667     .0050798
  dale | 0.0131947    .0349878    0.377   0.706     -.0553802    .0817696
  _cons | 0.349176     1.660689    0.210   0.833     -2.905714    3.604066

tau2  | _cons | 0.1246866   .0499158    2.498   0.012     .0268535    .2225198

ml model d0 d0progl0 (regco = poll over65 gnp gini educ dale) / tau2 if c20>0
  Number of obs = 21
  Wald chi2(6) = 11.28
  Log likelihood = -56.706088 Prob > chi2 = 0.0802

    | Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------|---------|--------------|------|--------|-----------------------------
  poll | 0.0010669    .025493    0.419   0.676     -.0039295    .0060634
  over65 | -0.0200338   .0488143   -0.410   0.682     -.115708    .0756404
  gnp | 0.0000873   .0000354    2.469   0.014     -.0000184    .0001566
  gini | 0.105159     .0144621    0.727   0.467     -.0178233    .0388611
  educ | -0.0145633   .0097144   -1.499   0.134     -.0336033    .0044766
  dale | -0.001246    .0318375   -0.039   0.969     -.0636463    .0611543
  _cons | 0.6808426    1.658054    0.411   0.681     -2.568883    3.930568

tau2  | _cons | 0.2300906    .0867875    2.651   0.008     .0599901    .400191
```

311
```
.ml model d0 d0prog21 (regco = poll over65 gnp gini educ dale) / tau2 if c21>0
Number of obs = 21
Wald chi2(6) = 21.63
Log likelihood = -41.592121
Prob > chi2 = 0.0014

|       | Coef.   | Std. Err. | z      | P>|z|    | [95% Conf. Interval] |
|-------|---------|-----------|--------|--------|-------------------|
| poll  | 0.0004138 | 0.0013287 | 0.311  | 0.755  | -0.0021905 to 0.0030181 |
| over65| -0.0163723 | 0.0310397 | -0.527 | 0.598  | -0.077209 to 0.044644 |
| gnp   | 3.17e-06 | 0.000194  | 0.163  | 0.870  | -0.0000349 to 0.0000413 |
| gini  | 0.000022  | 0.007641  | 0.003  | 0.998  | -0.014954 to 0.0149781 |
| educ  | -0.0205853 | 0.0056896 | -3.618 | 0.000  | -0.077209 to 0.044644 |
| dale  | 0.0500904 | 0.0212421 | 2.358  | 0.018  | 0.0084567 to 0.091724 |
| _cons | -1.282094 | 1.096724  | -1.169 | 0.242  | -3.431634 to 0.8674543 |

|       | Coef.   | Std. Err. | z      | P>|z|    | [95% Conf. Interval] |
|-------|---------|-----------|--------|--------|-------------------|
| poll  | -0.0006364 | 0.0025486 | -0.250 | 0.803  | -0.0056315 to 0.0043587 |
| over65| 0.0214666  | 0.055022  | 0.390  | 0.696  | -0.0863746 to 0.1293078 |
| gnp   | 0.000544  | 0.000308  | 1.765  | 0.078  | -6.02e-06 to 0.0001148 |
| gini  | 0.0176825  | 0.0163909 | 1.079  | 0.281  | -0.014432 to 0.0498081 |
| educ  | -0.0197888 | 0.0105649 | -1.873 | 0.061  | -0.040956 to 0.000918 |
| dale  | 0.0048797 | 0.0355903 | -0.154 | 0.877  | -0.075243 to 0.0642673 |
| _cons | 1.204919  | 1.845444  | 0.653  | 0.514  | -2.412084 to 4.821923 |

|       | Coef.   | Std. Err. | z      | P>|z|    | [95% Conf. Interval] |
|-------|---------|-----------|--------|--------|-------------------|
| poll  | -0.001356 | 0.0015867 | -0.855 | 0.393  | -0.0044658 to 0.0017539 |
| over65| 0.0680085 | 0.0376535 | -1.806 | 0.071  | -0.1418081 to 0.005791 |
| gnp   | 0.0002929 | 0.0002343 | 1.247  | 0.212  | -0.0000167 to 0.0000751 |
| gini  | 0.0056402 | 0.0118187 | 0.477  | 0.633  | -0.017524 to 0.0288045 |
| educ  | -0.0099035 | 0.0077477 | -1.278 | 0.201  | -0.0250888 to 0.0052819 |
| dale  | 0.0248575 | 0.0326064 | 0.762  | 0.446  | -0.0390498 to 0.0887649 |
| _cons | -0.0674224 | 1.511432  | -0.045 | 0.964  | -3.029776 to 2.894931 |
```

Number of obs = 21
Wald chi2(6) = 11.09
Log likelihood = -60.190957
Prob > chi2 = 0.0857

|       | Coef.   | Std. Err. | z      | P>|z|    | [95% Conf. Interval] |
|-------|---------|-----------|--------|--------|-------------------|
| poll  | -0.001356 | 0.0015867 | -0.855 | 0.393  | -0.0044658 to 0.0017539 |
| over65| 0.0680085 | 0.0376535 | -1.806 | 0.071  | -0.1418081 to 0.005791 |
| gnp   | 0.0002929 | 0.0002343 | 1.247  | 0.212  | -0.0000167 to 0.0000751 |
| gini  | 0.0056402 | 0.0118187 | 0.477  | 0.633  | -0.017524 to 0.0288045 |
| educ  | -0.0099035 | 0.0077477 | -1.278 | 0.201  | -0.0250888 to 0.0052819 |
| dale  | 0.0248575 | 0.0326064 | 0.762  | 0.446  | -0.0390498 to 0.0887649 |
| _cons | -0.0674224 | 1.511432  | -0.045 | 0.964  | -3.029776 to 2.894931 |
```

Number of obs = 21
Wald chi2(6) = 10.67
Log likelihood = -49.11387
Prob > chi2 = 0.0992
```
**. ml model d0 d0pro2 (regco = poll over65 gnp gini educ dale) / tau2 if c2>0**

Number of obs = 21  
Wald chi2(6) = 12.26  
Log likelihood = -78.018915  
Prob > chi2 = 0.0565

| regco | Coef. | Std. Err. | z     | P>|z|  | [95% Conf. Interval] |
|-------|-------|-----------|-------|------|---------------------|
| poll  | -0.006192 | 0.0026217 | -0.236 | 0.813 | -0.0057576, 0.0045192 |
| over65| -0.0106589 | 0.0494967 | -0.205 | 0.040 | -0.1986706, -0.0046472 |
| gnp   | 0.000117 | 0.0000667 | 1.754 | 0.079 | 0.0000138, 0.0002478 |
| dist  | -8.32e-07 | 1.46e-06 | -0.570 | 0.578 | -3.69e-06, 2.03e-06 |
| educ  | -0.023113 | 0.0100144 | -2.308 | 0.021 | -0.0427408, -0.0034852 |
| dale  | 0.0305877 | 0.0306722 | 0.997 | 0.321 | -0.0295287, 0.090704 |
| _cons | 0.453215 | 1.84617 | 0.245 | 0.806 | -3.165212, 4.071642 |

| tau2  | _cons | 0.2835818 | 0.1062127 | 2.670 | 0.008 | 0.0754087, 0.4917548 |

---

**. ml model d0 d0pro (regco = poll over65 gnp dist educ dale) / tau2 if c1>0**

Number of obs = 21  
Wald chi2(6) = 4.81  
Log likelihood = -60.497596  
Prob > chi2 = 0.5684

| regco | Coef. | Std. Err. | z     | P>|z|  | [95% Conf. Interval] |
|-------|-------|-----------|-------|------|---------------------|
| poll  | -0.005431 | 0.0016626 | -0.327 | 0.744 | -0.038017, 0.002755 |
| over65| -0.0278231 | 0.0316119 | -0.880 | 0.379 | -0.0897813, 0.0341351 |
| gnp   | -9.20e-06 | 0.0000369 | -0.249 | 0.803 | -0.0000815, 0.0000631 |
| dist  | 6.32e-07 | 1.00e-06 | 0.631 | 0.528 | -1.33e-06, 2.60e-06 |
| educ  | -0.0065429 | 0.0000667 | -1.087 | 0.277 | -0.0217576, 0.0689066 |
| dale  | 0.0235745 | 0.0061722 | 1.019 | 0.308 | -0.0000138, 0.0002478 |
| _cons | -0.3484194 | 1.341302 | -0.260 | 0.795 | -2.977324, 2.280485 |

| tau2  | _cons | 0.0766539 | 0.0310815 | 2.466 | 0.014 | 0.0157353, 0.1375724 |

---

**. ml model d0 d0prog (regco = poll over65 gnp gini educ dale) / tau2 if c2>0**

Number of obs = 21  
Wald chi2(6) = 10.94  
Log likelihood = -56.772328  
Prob > chi2 = 0.0904

| regco | Coef. | Std. Err. | z     | P>|z|  | [95% Conf. Interval] |
|-------|-------|-----------|-------|------|---------------------|
| poll  | 0.0024793 | 0.0026744 | 0.927 | 0.354 | 0.0000195, 0.000156 |
| over65| -0.0315979 | 0.0448144 | -0.705 | 0.481 | -0.1194325, 0.0562366 |
| gnp   | 0.0060877 | 0.000348 | 2.519 | 0.012 | 0.0000348, 0.000156 |
| gini  | 0.001463 | 0.014192 | 0.292 | 0.770 | -0.0236695, 0.0319621 |
| educ  | 0.0130661 | 0.0314044 | 0.416 | 0.677 | -0.0484854, 0.0746177 |
| _cons | 0.2302503 | 1.662984 | 0.138 | 0.890 | -3.029138, 3.489638 |

| tau2  | _cons | 0.2231022 | 0.0862486 | 2.587 | 0.010 | 0.054058, 0.3921464 |

---

313
. ml model d0 d0pro3 (regco = poll over65 gnpj dist educ dale) / tau2 if c3>0
Number of obs = 21
Wald chi2(6) = 11.96
Prob > chi2 = 0.0630
Log likelihood = -65.572493

| Coef.  | Std. Err. | z     | P>|z|  | [95% Conf. Interval] |
|--------|-----------|-------|------|----------------------------|
| poll   | -.0000771 | .0013923 | -0.055 | 0.956 | -.002086 -.0026518 |
| over65 | -.0204152 | .0340243 | -0.600 | 0.548 | -.0871016 .0462713 |
| gnp    | -.0000154 | .0000388 | -0.398 | 0.691 | -.0000914 .0000605 |
| dist   | 7.86e-07  | 9.00e-07 | 0.874 | 0.382 | -9.77e-07 2.55e-06 |
| educ   | -.0190203 | .0065926 | -2.885 | 0.004 | -.0319415 -.0060991 |
| dale   | .0418641  | .0231024 | 1.812 | 0.070 | -.0034519 .087144  |
| _cons  | -.7634688 | 1.296464 | -0.589 | 0.556 | -3.044911 1.777554 |

tau2
| _cons  | .0696536 | .0378484 | 1.840 | 0.066 | -.0045279 .1438352 |

. ml model d0 d0pro4 (regco = poll over65 gnpj dist educ dale) / tau2 if c4>0
Number of obs = 21
Wald chi2(6) = 11.37
Log likelihood = -75.447562

| Coef.  | Std. Err. | z     | P>|z|  | [95% Conf. Interval] |
|--------|-----------|-------|------|----------------------------|
| poll   | .003088   | .0354656 | 0.079 | 0.939 | -.0166334 -.0030375 |
| over65 | .0183e-06 | 1.48e-06 | 0.021 | 0.982 | -.001003 .001255  |
| gnp    | -.0190203 | .0084025 | -1.980 | 0.048 | -.0331019 -.001648 |
| dist   | .0418641  | .0231024 | 1.812 | 0.070 | -.0034519 .087144  |
| educ   | -.030375  | .0128789 | -0.105 | 0.916 | .059639 .0535641  |
| dale   | .7467562  | 1.709996 | 0.437 | 0.662 | -2.604775 4.098287 |
| _cons  | .2300362  | .0875023 | 2.629 | 0.009 | .0585355 .4015369 |

tau2
| _cons  | .2300362 | .0875023 | 2.629 | 0.009 | .0585355 .4015369 |

. ml model d0 d0pro5 (regco = poll over65 gnpj dist educ dale) / tau2 if c5>0
Number of obs = 21
Wald chi2(6) = 22.62
Log likelihood = -58.125198

| Coef.  | Std. Err. | z     | P>|z|  | [95% Conf. Interval] |
|--------|-----------|-------|------|----------------------------|
| poll   | .0004273  | .0014068 | 0.304 | 0.761 | -.0023299 .0031845 |
| over65 | -.1096787 | .0256529 | -4.275 | 0.000 | -.1599574 -.0594  |
| gnp    | .0000739  | .0000319 | 2.312 | 0.021 | .0000113 .0001365 |
| dist   | -.821e-07 | 7.08e-07 | -1.159 | 0.247 | -2.21e-06 5.67e-07 |
| educ   | -.0078769 | .0056303 | -1.399 | 0.162 | -.018912 .0031583 |
| dale   | .0362322  | .0191027 | 1.897 | 0.058 | -.0012085 .0736729 |
| _cons  | -.8023469 | 1.1521 | -0.696 | 0.486 | -3.060421 1.455727 |

tau2
| _cons  | .0512909 | .0240112 | 2.136 | 0.033 | .0042299 .098352 |

314
```
. ml model d0 d0pro6 (regco = poll over65 gnp dist educ dale) / tau2 if c6>0
Number of obs  =  21
Wald chi2(6)   =  9.76
Log likelihood = -78.58854  Prob > chi2  =  0.1352

+------------------------------------------------------------+
|               | Coef.  | Std. Err. |      z | P>|z|   | [95% Conf. Interval]                          |
|------------------------------------------------------------+
| poll                | -0.0006427 | 0.0029719 | -0.216 | 0.829 | -0.0064675 to 0.0051822                     |
| over65              | -0.0011285 | 0.056405  | -0.020 | 0.984 | -0.1116803 to 0.1094232                    |
| gnp                 | -0.000152  | 0.000677  | -0.225 | 0.822 | -0.000148 to 0.0001175                     |
| dist                | 1.92e-06   | 1.69e-06  | 1.137  | 0.255 | -1.39e-06 to 5.23e-06                      |
| educ                | -0.0246351 | 0.0104321 | -2.361 | 0.020 | -0.0450816 to 0.0041885                    |
| dale                | 0.0214091  | 0.0331944 | 0.645  | 0.519 | -0.0436507 to 0.0864688                    |
| _cons               | 0.6189596  | 1.941824  | 0.319  | 0.750 | -3.186945 to 4.424864                      |
+------------------------------------------------------------+

tau2 | _cons | .3241059 | 1.150498 | 2.817   | 0.005   | 0.0986124 to .5495994                     |
+------------------------------------------------------------+

. ml model d0 d0pro7 (regco = poll over65 gnp dist educ dale) / tau2 if c7>0
Number of obs  =  21
Wald chi2(6)   =  10.55
Log likelihood = -69.649745  Prob > chi2  =  0.1034

+------------------------------------------------------------+
|               | Coef.  | Std. Err. |      z | P>|z|   | [95% Conf. Interval]                          |
|------------------------------------------------------------+
| poll                | -0.0001605 | 0.0017096 | -0.094 | 0.925 | -0.0035112 to 0.0031903                    |
| over65              | -0.0413539 | 0.0393238 | -1.052 | 0.293 | -0.1184271 to 0.0357193                    |
| gnp                 | -0.000014  | 0.0000462 | -0.303 | 0.762 | -0.0001046 to 0.0000766                    |
| dist                | 1.67e-06   | 1.09e-06  | 1.533  | 0.125 | -4.65e-07 to 3.81e-06                      |
| educ                | -0.0105373 | 0.0074536 | -1.414 | 0.157 | -0.0251461 to 0.0040715                   |
| dale                | 0.016358   | 0.028497  | 0.574  | 0.566 | -0.0394951 to 0.0722111                   |
| _cons               | 0.2807918  | 1.609129  | 0.174  | 0.861 | -2.873044 to 3.434627                     |
+------------------------------------------------------------+

tau2 | _cons | .1247839 | .0491282 | 2.540   | 0.011   | 0.0284944 to .2210734                     |
+------------------------------------------------------------+

. ml model d0 d0pro8 (regco = poll over65 gnp dist educ dale) / tau2 if c8>0
Number of obs  =  21
Wald chi2(6)   =  11.55
Log likelihood = -76.281458  Prob > chi2  =  0.0729

+------------------------------------------------------------+
|               | Coef.  | Std. Err. |      z | P>|z|   | [95% Conf. Interval]                          |
|------------------------------------------------------------+
| poll                | 0.0021756 | 0.0026497 | 0.821  | 0.412 | -0.0030176 to 0.0073688                     |
| over65              | -0.0600763 | 0.0457228 | -1.314 | 0.189 | -0.1496913 to 0.0295387                    |
| gnp                 | 0.000973  | 0.000624  | 1.559  | 0.119 | -0.00025 to 0.0002195                     |
| dist                | -2.94e-08 | 1.39e-06  | -2.021 | 0.045 | -2.76e-06 to 2.70e-06                      |
| educ                | -0.0150989 | 0.0093288 | -1.619 | 0.106 | -0.0333829 to 0.0031852                    |
| dale                | 0.0179741  | 0.0278018 | 0.647  | 0.518 | -0.0365164 to 0.0724646                    |
| _cons               | -0.0228719 | 1.693375  | -0.014 | 0.989 | -3.341826 to 3.296083                      |
+------------------------------------------------------------+

tau2 | _cons | .2315559 | .0914157 | 2.533   | 0.011   | 0.0523845 to .4107273                     |
+------------------------------------------------------------+
```
```
. ml model d0 d0pro9 (regco = poll over65 gnp dist educ dale) / tau2 if c9>0
Number of obs = 21
Wald chi2(6) = 20.25
Log likelihood = -55.524752 Prob > chi2 = 0.0025

| regco   | Coef.   | Std. Err. | z     | P>|z| | [95% Conf. Interval] |
|---------|---------|-----------|-------|-----|---------------------|
| poll    | .0006483 | .0011288  | 0.574 | 0.566 | -.0015641 .0028607 |
| over65  | -.0373089 | .0245416  | -1.520 | 0.128 | -.0854096 .0107918 |
| gnp     | .0000179  | .0000273  | 0.656 | 0.512 | -.0000355 .0000713 |
| dist    | 3.58e-08  | 5.97e-07  | 0.600 | 0.952 | -1.13e-06 1.21e-06 |
| educ    | -.0171846 | .0051096  | -3.363 | 0.001 | -.0271992 .000717  |
| dale    | .0427803  | .01620999 | 2.639 | 0.008 | .0110094 .0745512  |
| _cons   | -1.026349 | .9723321  | -1.056 | 0.291 | -2.932085 .8793869 |

tau2
| _cons   | .0267869 | .01464    | 1.830 | 0.067 | -.001907 .0554809 |

. ml model d0 d0pro10 (regco = poll over65 gnp dist educ dale) / tau2 if c10>0
Number of obs = 21
Wald chi2(6) = 10.30
Log likelihood = -78.306758 Prob > chi2 = 0.1124

| regco   | Coef.   | Std. Err. | z     | P>|z| | [95% Conf. Interval] |
|---------|---------|-----------|-------|-----|---------------------|
| poll    | -.0004811 | .0026426  | -0.182 | 0.856 | -.0056605 .0046983 |
| over65  | -.006102  | .0500868  | -0.122 | 0.903 | -.1042704 .0920663 |
| gnp     | .000035   | .0000664  | 0.527 | 0.598 | -.0000951 .0001651 |
| dist    | 7.60e-07  | 1.50e-06  | 0.506 | 0.613 | -2.18e-06 3.70e-06 |
| educ    | -.0238628 | .0097586  | -2.445 | 0.014 | -.0429893 .0047362 |
| dale    | .011543   | .0309565  | 0.372 | 0.710 | -.0491594 .072188  |
| _cons   | 1.186873  | 1.840355  | 0.645 | 0.519 | -2.420156 4.793902 |

tau2
| _cons   | .2785972 | .106044   | 2.627 | 0.009 | .0707547 .4864396 |

. ml model d0 d0pro11 (regco = poll over65 gnp dist educ dale) / tau2 if c11>0
Number of obs = 21
Wald chi2(6) = 14.81
Log likelihood = -65.109379 Prob > chi2 = 0.0218

| regco   | Coef.   | Std. Err. | z     | P>|z| | [95% Conf. Interval] |
|---------|---------|-----------|-------|-----|---------------------|
| poll    | -.0005965 | .0014003  | -0.426 | 0.670 | -.0033409 .0021479 |
| over65  | -.0993418 | .0304978  | -3.257 | 0.001 | -.1591165 -.0395672 |
| gnp     | .000449   | .0003631  | 1.242 | 0.214 | -.000026 .0001157 |
| dist    | -2.93e-07 | 8.98e-07  | -0.326 | 0.744 | -2.05e-06 1.47e-06 |
| educ    | -.0102513 | .0063978  | -1.602 | 0.109 | -.0227909 .0022882 |
| dale    | .0418382  | .0224232  | 1.866 | 0.062 | -.0021105 .0857868 |
| _cons   | -.817001  | 1.259031  | -0.649 | 0.516 | -3.284656 1.650653 |

tau2
| _cons   | .066666  | .0336532  | 1.981 | 0.048 | .0007069 .1326251 |

316```
## Model 1: d0pro12

```
.ml model d0 d0pro12 (regco = poll over65 gnp dist educ dale) / tau2 if cl2>0
Number of obs = 21
Wald chi2(6) = 12.89
Log likelihood = -79.246571 Prob > chi2 = 0.0449

|       | Coef.     | Std. Err. | z     | P>|z|    | [95% Conf. Interval] |
|-------|-----------|-----------|-------|--------|----------------------|
| poll  | .0025829  | .0027126  | 0.952 | 0.341  | -.0027336 -.0078994 |
| over65| .0050301  | .048515   | 0.104 | 0.917  | -.0900575 .101177  |
| gnp   | .0000486  | .0000628  | 0.774 | 0.439  | -.0000745 .001717  |
| dist  | 8.53e-07  | 1.47e-06  | 0.580 | 0.562  | -2.03e-06 3.74e-06 |
| educ  | -.0249288 | .0094715  | 0.263 | 0.794  | -.0900575 .101177  |
| dale  | .0179483  | .0281221  | 0.638 | 0.523  | -.03717 .0730667  |
| _cons | .1861631  | 1.686524  | 0.110 | 0.912  | -3.119363 3.491689 |
```

## Model 2: d0pro13

```
.ml model d0 d0pro13 (regco = poll over65 gnp dist educ dale) / tau2 if cl3>0
Number of obs = 21
Wald chi2(6) = 5.88
Log likelihood = -72.58448 Prob > chi2 = 0.4373

|       | Coef.     | Std. Err. | z     | P>|z|    | [95% Conf. Interval] |
|-------|-----------|-----------|-------|--------|----------------------|
| poll  | -.0005516 | .0019391  | -0.284| 0.776  | -.0043521 .0032489 |
| over65| -.0523958 | .0356259  | -1.471| 0.141  | -.1222213 .0174297 |
| gnp   | 6.65e-06  | .000042   | 0.158 | 0.874  | -.0000756 .0000889 |
| dist  | 5.67e-07  | 1.14e-06  | 0.495 | 0.620  | -1.68e-06 2.81e-06 |
| educ  | -.0070712 | .0066982  | -1.056| 0.291  | -.0201995 .0060571 |
| dale  | .0186084  | .0272949  | 0.682 | 0.495  | -.0348887 .0721055 |
| _cons | .0953264  | 1.575951  | 0.060 | 0.952  | -2.99348 3.184133 |
```

## Model 3: d0pro14

```
.ml model d0 d0pro14 (regco = poll over65 gnp dist educ dale) / tau2 if cl4>0
Number of obs = 21
Wald chi2(6) = 10.17
Log likelihood = -76.868233 Prob > chi2 = 0.1177

|       | Coef.     | Std. Err. | z     | P>|z|    | [95% Conf. Interval] |
|-------|-----------|-----------|-------|--------|----------------------|
| poll  | .0003532  | .002826   | 0.125 | 0.901  | -.0051856 .005892  |
| over65| -.0728673 | .0500353  | -1.456| 0.145  | -.1709348 .0252001 |
| gnp   | .0000774  | .0000629  | 1.230 | 0.219  | -.000046 .0002008 |
| dist  | 3.28e-07  | 1.49e-06  | 0.220 | 0.826  | -2.60e-06 3.25e-06 |
| educ  | -.0193666 | .0095437  | -2.029| 0.042  | -.0380719 -.0006614|
| dale  | .0234045  | .0301878  | 0.842 | 0.400  | -.0337625 .0845716 |
| _cons | .2222238  | 1.797065  | 0.124 | 0.902  | -3.299959 3.744406 |
```

## tau2

```
<table>
<thead>
<tr>
<th>_cons</th>
<th>0.2271802</th>
<th>0.0945557</th>
<th>2.403</th>
<th>0.016</th>
<th>0.0418544 .4125059</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>tau2</th>
</tr>
</thead>
<tbody>
<tr>
<td>_cons</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>tau2</th>
</tr>
</thead>
<tbody>
<tr>
<td>_cons</td>
</tr>
</tbody>
</table>
```
```
Number of obs = 21
Wald chi2(6) = 15.47
Log likelihood = -60.680402 Prob > chi2 = 0.0169

|          | Coef.  | Std. Err. | z     | P>|z|  | [95% Conf. Interval]     |
|----------|--------|-----------|-------|-----|--------------------------|
| polll    | 0.0002821 | 0.0011848 | -0.98  | 0.32 | -1.204 0.229 | -0.0943155 0.025298 |
| over65   | -0.0358928 | 0.029808 | -1.20 | 0.23 | -0.007 0.995 | -0.0000643 0.0000639 |
| gnp      | -2.20e-07  | 0.000327 | -0.007 | 0.995 | -0.0000643 0.0000639 |
| dist     | 4.25e-07  | 7.72e-07 | 0.551 | 0.581 | -1.09e-06 1.94e-06 |
| educ     | -0.019984 | 0.0061018 | -3.275 | 0.001 | -0.0319434 -0.0080246 |
| dale     | 0.0482014 | 0.0201695 | 2.390 | 0.017 | 0.0061018 0.0201695 |
| _cons    | 1.009914 | 1.130495 | -0.893 | 0.372 | -1.225644 1.205815 |

Number of obs = 21
Wald chi2(6) = 10.87
Log likelihood = -77.156455 Prob > chi2 = 0.0924

|          | Coef.  | Std. Err. | z     | P>|z|  | [95% Conf. Interval]     |
|----------|--------|-----------|-------|-----|--------------------------|
| polll    | 0.0030332 | 0.0032268 | 0.0003  | 0.999 | 0.0003 0.0003 | -0.0164941 -0.0086707 |
| over65   | 0.0000229 | 0.0000638 | 0.0359 | 0.719 | -0.0001022 0.0001481 |
| gnp      | 2.07e-06  | 1.62e-06  | 1.278 | 0.201 | -1.11e-06 5.25e-06 |
| dist     | -0.0164941 | 0.0094415 | -1.747 | 0.081 | -0.0349991 0.032011 |
| educ     | -0.0086707 | 0.0316401 | -0.274 | 0.784 | -0.0706842 0.053328 |
| dale     | 1.207074  | 1.875864  | 0.643 | 0.520 | -2.469351 4.883699 |
| _cons    | 0.2869027 | 0.1010228 | 2.840 | 0.005 | 0.0889017 0.4849037 |

Number of obs = 21
Wald chi2(6) = 4.40
Log likelihood = -64.621635 Prob > chi2 = 0.6232

|          | Coef.  | Std. Err. | z     | P>|z|  | [95% Conf. Interval]     |
|----------|--------|-----------|-------|-----|--------------------------|
| polll    | -0.0008413 | 0.0018175 | -4.63 | 0.643 | -0.0044037 0.002721 |
| over65   | -0.0346601 | 0.0338917 | -1.023 | 0.306 | -1.010866 0.317664 |
| gnp      | 3.18e-06  | 0.000643  | 0.075 | 0.940 | -0.0000797 0.000086 |
| dist     | 4.11e-07  | 1.05e-06  | 0.392 | 0.695 | -1.65e-06 2.47e-06 |
| educ     | -0.0057568 | 0.0067782 | -0.849 | 0.396 | -0.019042 0.0075283 |
| dale     | 0.0211657 | 0.0253077 | 0.836 | 0.403 | -0.0284365 0.0707678 |
| _cons    | -1.1742026 | 1.453949 | -0.120 | 0.905 | -3.023891 2.675486 |
```

. ml model d0 d0pro18 (regco = poll over65 gnp dist educ dale ) / tau2 if cl8>0

Number of obs = 21
Wald chi2(6) = 12.59
Log likelihood = -77.387389
Prob > chi2 = 0.0500

| regco  | Coef.   | Std. Err. | z  | P>|z|  | [95% Conf. Interval] |
|--------|---------|-----------|----|------|-------------------|
| poll   | 0.0002704 | 0.0028722 | 0.094 | 0.925 | -0.005359 - 0.0058999 |
| over65 | 0.0038843 | 0.0524899 | 0.074 | 0.941 | -0.0989941 - 0.1067626 |
| gnp    | 0.0000223 | 0.0006484 | 0.344 | 0.731 | -0.0001046 - 0.0001492 |
| dist   | 1.23e-06  | 1.50e-06  | 0.822 | 0.411 | -1.71e-06 - 4.18e-06 |
| educ   | -0.029914 | 0.098927  | -3.024 | 0.002 | -0.493032 - 0.0105247 |
| dale   | 0.0237677 | 0.0305867 | 0.777 | 0.437 | -0.0361829 - 0.0837182 |
| _cons  | 0.5886153 | 1.828479  | 0.322 | 0.748 | -2.995137 - 4.172368 |

Number of obs = 21
Wald chi2(6) = 11.46
Log likelihood = -67.19496
Prob > chi2 = 0.0751

| regco  | Coef.   | Std. Err. | z  | P>|z|  | [95% Conf. Interval] |
|--------|---------|-----------|----|------|-------------------|
| poll   | -0.0005875 | 0.0017487 | -0.336 | 0.737 | -0.0040149 - 0.00284 |
| over65 | -0.0782362 | 0.037872 | -2.066 | 0.039 | -0.1524639 - 0.040885 |
| gnp    | 0.0000215  | 0.0004444 | 0.484 | 0.628 | -0.0000655 - 0.0001086 |
| dist   | 6.35e-07   | 1.15e-06  | 0.554 | 0.580 | -1.61e-06 - 2.88e-06 |
| educ   | -0.0133911 | 0.0073358 | -1.825 | 0.068 | -0.027769 - 0.009868 |
| dale   | 0.0274538  | 0.029016  | 0.944 | 0.345 | -0.0295451 - 0.0844526 |
| _cons  | 0.0974354  | 1.626762  | 0.060 | 0.952 | -3.090959 - 3.28583 |

Number of obs = 21
Wald chi2(6) = 10.88
Log likelihood = -75.208808
Prob > chi2 = 0.0922

| regco  | Coef.   | Std. Err. | z  | P>|z|  | [95% Conf. Interval] |
|--------|---------|-----------|----|------|-------------------|
| poll   | 0.0012632 | 0.0026155 | 0.483 | 0.629 | -0.0038631 - 0.0063894 |
| over65 | 0.0289927 | 0.0457436 | -0.634 | 0.526 | -0.1186485 - 0.0606631 |
| gnp    | 0.000063  | 0.000616  | 1.022 | 0.307 | -0.0000577 - 0.0001837 |
| dist   | 8.05e-07  | 1.49e-06  | 0.539 | 0.590 | -2.12e-06 - 3.73e-06 |
| educ   | -0.0163646 | 0.0091299 | -1.780 | 0.075 | -0.0343824 - 0.016532 |
| dale   | 0.0064004  | 0.0287015 | 0.223 | 0.823 | -0.0498499 - 0.0626579 |
| _cons  | 0.7199767  | 1.685094  | 0.427 | 0.669 | -2.582746 - 4.0227 |

Number of obs = 21
Wald chi2(6) = 12.59
Log likelihood = -77.387389
Prob > chi2 = 0.0500
. ml model d0 d0pro21 (regco = poll over65 gnp dist educ dale) / tau2 if c21>0
   Number of obs = 21
   Wald chi2(6) = 22.75
   Log likelihood = -59.812367  Prob > chi2 = 0.0009

| regco | Coef. | Std. Err. | z     | P>|z| | [95% Conf. Interval] |
|-------|-------|-----------|-------|------|----------------------|
| poll  | 0.0001983 | 0.0013292 | 0.149 | 0.881 | -0.002407 .0028035 |
| over65| -0.0266413 | 0.0272508 | -0.978 | 0.328 | -0.080052 .0267693 |
| gnp   | 0.0000232  | 0.0000298 | 0.777 | 0.437 | -0.0000353 .0000817 |
| dist  | -5.94e-07  | 6.93e-07  | -0.857 | 0.391 | -1.95e-06 7.65e-07 |
| educ  | -0.0208296 | 0.054341  | -3.833 | 0.000 | -0.0314802 .010179 |
| dale  | 0.0535833  | 0.0179491 | 2.985 | 0.003 | 0.0184035 .0887631 |
| _cons | -1.372144  | 1.061943  | -1.292 | 0.196 | -3.453514 .709226 |

| tau2  | _cons   | 0.0379386 | 0.0208807 | 1.817 | 0.069 | -0.0029867 .078864 |

. ml model d0 d0pro22 (regco = poll over65 gnp dist educ dale) / tau2 if c22>0
   Number of obs = 21
   Wald chi2(6) = 9.97
   Log likelihood = -79.175377  Prob > chi2 = 0.1259

| regco | Coef. | Std. Err. | z     | P>|z| | [95% Conf. Interval] |
|-------|-------|-----------|-------|------|----------------------|
| poll  | -0.0006727 | 0.0026079 | -0.258 | 0.796 | -0.0057842 .0044387 |
| over65| -0.0002732 | 0.0512387 | -0.005 | 0.996 | -1.006996 .010132 |
| gnp   | 0.0000286  | 0.0000688 | 0.415 | 0.678 | -0.0001062 .0001634 |
| dist  | 8.58e-07   | 1.56e-06  | 0.550 | 0.582 | -2.20e-06 3.91e-06 |
| educ  | -0.0235602 | 0.0100124 | -2.353 | 0.019 | -0.0431841 -.0039362 |
| dale  | 0.0106132  | 0.018721  | 0.333 | 0.739 | -0.051855 .0730814 |
| _cons | 1.223484   | 1.89204   | 0.647 | 0.518 | -2.484846 4.931814 |

| tau2  | _cons   | 0.293957 | 0.1117149 | 2.631 | 0.009 | 0.0749997 .5129142 |

. ml model d0 d0pro23 (regco = poll over65 gnp dist educ dale) / tau2 if c23>0
   Number of obs = 21
   Wald chi2(6) = 10.66
   Log likelihood = -67.786131  Prob > chi2 = 0.0995

| regco | Coef. | Std. Err. | z     | P>|z| | [95% Conf. Interval] |
|-------|-------|-----------|-------|------|----------------------|
| poll  | -0.0012702 | 0.0015677 | -0.810 | 0.418 | -0.0043429 .0018025 |
| over65| -0.0799726 | 0.0324277 | -2.466 | 0.014 | -1.1435298 .0164154 |
| gnp   | 0.0000315  | 0.0000409 | 0.771 | 0.441 | -0.0000486 .0001117 |
| dist  | -8.47e-08  | 1.09e-06  | -0.077 | 0.938 | -2.23e-06 2.06e-06 |
| educ  | -0.012043  | 0.0067827 | -1.776 | 0.076 | -0.0253369 .001251 |
| dale  | 0.0357296  | 0.0263827 | 1.354 | 0.176 | -0.0159796 .0874388 |
| _cons | -3.160896  | 1.456501  | -0.217 | 0.828 | -3.170778 2.538599 |

| tau2  | _cons   | 0.0962779 | 0.044003 | 2.188 | 0.029 | 0.0100336 .1825222 |

320
```
.ml model d0 d0pro24 (regco = poll over65 gnp dist educ dale) / tau2 if c24>0.
Number of obs = 21
Wald chi2(6) = 10.80
Prob > chi2 = 0.0946

Log likelihood = -75.175338

|      | Coef.    | Std. Err. |       z |     P>|z|     |   [95% Conf. Interval] |
|------|----------|-----------|---------|--------|------------------------|
| regco|          |           |         |        |                        |
| poll | 0.0025545| 0.0027103 | 0.942   | 0.346  | -0.0027577 .0078666   |
| over65| -0.0359248| 0.0415632 | -0.864 | 0.387  | -0.1173871 .0455376   |
| gnp  | 0.0000816| 0.0000603 | 1.353   | 0.176  | -0.0000366 .0001999   |
| dist | 2.19e-07 | 1.46e-06  | 0.149   | 0.881  | -2.65e-06 3.09e-06    |
| educ | -0.0204541| 0.0090793 | -2.253  | 0.024  | -0.0382491 .002659    |
| dale | 0.0166926 | 0.0280277 | 0.596   | 0.551  | -0.0382407 .0716259   |
| _cons| 0.2262885 | 1.675956  | 0.135   | 0.893  | -3.058524 3.511101    |

|      | Coef.    | Std. Err. |       z |     P>|z|     |   [95% Conf. Interval] |
|------|----------|-----------|---------|--------|------------------------|
| tau2 |          |           |         |        |                        |
| _cons| 0.2254896| 0.0866677 | 2.602   | 0.009  | 0.055624 .3953552     |
```