PRICE INDICES, MONETARY ANALYSIS, AND INFLATION

A MACRO-THEORETICAL INVESTIGATION

Sergio Rossi

University College London
Department of Economics

Ph.D. Thesis
2000
ABSTRACT

This thesis investigates inflation from a macro-theoretical point of view. It originates in a critical appraisal of traditional inflation analysis, where the latter phenomenon is identified with an ongoing increase in the level of aggregate prices because 'too much money is chasing too few goods'. It argues that the prevalent idea of money and output being two separate and autonomous objects can neither explain the value of money nor its variations over time. It also argues that output as a whole cannot be measured in this widely-shared analytical framework. A new theory is called for. The problem of inflation in the alternative framework developed in the thesis reveals its essentially macroeconomic nature. It is argued that to gain an understanding of inflation it is necessary to focus analysis on the formation of national income and not on its distribution. Within the proposed new framework, the production process is investigated in terms of flows, rather than in terms of stocks changing hands in the process of output circulation; both money and banking are instrumental in generating, and measuring, macroeconomic magnitudes. Elaboration of the role of money and banking in this analytical framework provides a number of theoretical propositions that lead to the conclusion that the origin of inflation is 'monetary' and 'structural' rather than 'real' and 'behavioural'. The monetary-structural element is the design of the payment system; when this system is redesigned to recognise the macroeconomic relationship existing between bank money and production, inflationary pressures are at least revealed and may be eliminated. This thesis seeks to contribute to such a conceptual design.
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>2</td>
</tr>
<tr>
<td>CONTENTS</td>
<td>3</td>
</tr>
<tr>
<td>TABLES</td>
<td>6</td>
</tr>
<tr>
<td>FIGURES</td>
<td>7</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>8</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>9</td>
</tr>
</tbody>
</table>

### Part I Methodological issues in the measurement of inflation

1. **THE METHODOLOGICAL DEBATE IN TRADITIONAL ANALYSIS**
   - Some notes on the conventional measures of inflation
     - Defining inflation                                           | 20   |
     - Introducing price index analysis                            | 20   |
     - Selecting the appropriate formula                           | 25   |
   - Debated methodological issues in traditional analysis        | 35   |
     - The compositional issue of the representative market basket | 39   |
     - From consumer price indices to cost-of-living indices       | 47   |

2. **FROM TECHNICAL BIASES TO ANALYTICAL ISSUES**
   - Measurement problems in aggregate price indices              | 54   |
     - The substitution bias                                       | 54   |
     - The new-goods bias                                           | 59   |
     - The quality-change bias                                      | 62   |
   - Analytical causes of inflation measurement problems          | 71   |
     - The representative market basket and total current supply   | 71   |
     - The sample of population and total current demand           | 77   |
     - The problem of aggregating microeconomic data               | 80   |
Part II Towards a macroeconomic analysis of inflationary disequilibria

3 THE NEOCLASSICAL ANALYSIS OF INFLATION: A CRITICAL APPRAISAL 91
   The optimum quantity of money and other issues 94
   The Fisherine quantity equation 94
   The Cambridge quantity equation 101
   The modern quantity equation and the expected inflation rate 104
   The dichotomous representation of monetary economies 110
   Money and output as two distinct objects 112
   Money, a veil? 115
   Homogeneity postulate and inflation 119

4 THE ARGUMENT REFINED: EXOGENOUS AND ENDOGENOUS MONEY 123
   Exogenous money, or the dichotomy revisited 124
   The medium of exchange, a monetary asset? 126
   The Cartalist idea of outside money and the inflation tax 135
   Endogenous money, or the association of money and output 146
   Money, a numéraire? 146
   The means of payment, a double-entry integer 150

Part III A modern paradigm for inflation analysis

5 WAGE SETTING, CREDIT POLICY, AND INFLATION 160
   The macroeconomics of the wage bargain 161
   Labour productivity, effective demand, and money wages 161
   The conflict inflation approach: a critical appraisal 177
   Bank credit and inflation 187
   Excess credit facilities and inflation: a reconsideration 188
   Inflation: is there any macroeconomic harm? 198

6 INFLATION AND CAPITAL ACCUMULATION 209
   Empty money and inflation 211
   Empty money, forced saving, and fixed capital 212
   The amortisation of fixed capital 216
   An outline of structural monetary reform 222
   On the book-entry distinction between money and credit 223
   On the necessary introduction of a fixed capital department 234
7 MAIN CONCLUSIONS AND PERSPECTIVES FOR FURTHER INQUIRY
   Aim of the thesis and methodology chosen 244
   A methodological and heuristic assessment of our analysis 244
   Criticisms raised in the interpretation of Schmitt's theory 258
   Some perspectives for further research 264
   Epilogue 267

BIBLIOGRAPHY 269
TABLES

Table 1.1  A typology of inflationary phenomena 23
Table 1.2  Frequency of weighting revisions of some official price indices 36
Table 4.1  The inflation tax: real goods vs. nominal money 140
Table 4.2  The financing of public sector deficits 145
Table 4.3  The result of an absolute exchange on the factor market 152
Table 5.1  A speculation-led excess of bank credit 190
Table 5.2  The macroeconomic result of excess credit facilities 192
Table 6.1  The bi-partition of the banks' bookkeeping 225
Table 6.2  Savings, the financial market, and the two departments 228
Table 6.3  The working of the fixed capital department 238
FIGURES

Figure 1.1 The concept of price level 24
Figure 2.1 The macroeconomic inadequacy of the representative basket 74
Figure 2.2 Total demand and representative demand 79
Figure 3.1 The concept of relative exchange 95
Figure 4.1 The medium-of-exchange function of 'monetary assets' 126
Figure 4.2 The sale of national output in neoclassical analysis 132
Figure 4.3 The exchange of fiat money for goods and services 141
Figure 4.4 The two absolute exchanges of national output 155
Figure 5.1 Inclusion of a positive expenditure into a negative expenditure 204
ACKNOWLEDGEMENTS

It is the author's pleasure to acknowledge his intellectual debt to a number of people who have contributed to the development of this work in several important ways. In this regard, I would like to thank wholeheartedly my two Supervisors: Professors Victoria Chick and Lord Desai. Though they would not agree entirely with the conceptual framework that shapes the diagnosis developed in this thesis, their constant support and critical contribution have been an invaluable help in the completion of this research. Without implicating them for any remaining error or opinions expressed here, they should also be warmly thanked for the privilege they granted to me to work so closely with them and thus to benefit from their world-known expertise. I am also most grateful to Professors Heinrich Bortis, Alvaro Cencini, Bernard Dafflon, and Bernard Schmitt for their advice, encouragement and constructive criticism over the whole period I have spent in the UK since my first post-doctoral research year at the London School of Economics. I would also like to thank many people and institutions for the provision of scientific work that would have been otherwise impossible to obtain; in particular, the Bank of England, the Deutsche Bundesbank, the European Central Bank, and the Swiss Federal Statistical Office (especially Doctors Lorenzo Cascioni and Marcello Corti). Simona Cain made an important contribution in improving the style of Chapters 1 to 4. I am indebted to her as well as to the Swiss National Science Foundation, the Divizione della cultura del Cantone Ticino, the Fondazione Arturo & Marguerite Lang, and the Fondazione Felix Leemann for their financial support. Although the author is the sole responsible for the result, this work has also been possible thanks to an Overseas Research Students Award by the Committee of Vice-Chancellors and Principals of the United Kingdom, which is gratefully acknowledged. At a more personal level, I am most grateful to my mother for her constant, and multifaceted, support, especially when stress and discouragement have put the completion of this research at risk. This thesis is dedicated to her.
INTRODUCTION

This thesis aims at three goals.

The first is to discuss the methodological issues involved in the measurement of inflation, as established by reference to index number theory and its specific application, that is, the retail (or consumer) price index. As the influential Chairman of the Board of Governors of the US Federal Reserve System has recently observed, '[t]he remarkable progress that has been made by virtually all of the major industrial countries in achieving low rates of inflation in recent years has brought into sharper focus the issue of price measurement' (Greenspan 1997b: 1). Granting this progress, a new set of issues is now emerging on the agenda of economic policy makers. 'As we move closer to price stability, the necessity of measuring prices accurately has become an especial challenge. Biases of a few tenths in annual inflation rates do not matter when inflation is high. They do matter when, as now, a debate has emerged over whether our economies are moving toward price deflation' (p. 1). Accurately measuring prices and the rate of change of aggregate price levels is indeed of central importance for current investigation into our market-based monetary economies of production. It is in particular necessary for analysing economic developments, including output growth, government spending, and poverty rates, as well as for conducting macroeconomic policy. In Greenspan's own words, '[i]f the general price level is estimated to be rising more rapidly than is in fact the case, then we are simultaneously understating growth in real GDP and productivity, and real incomes and living standards are rising faster than our published data suggest' (Greenspan 1998: 1)\(^1\). As a matter of fact, despite the advances in price measurement that have

\(^1\) See also Blow and Crawford (1999b: 1). The implications of overstating price increases for the general government budgetary policy are equally important. In the United States, for instance, it has been calculated that an overestimation of price increases of 1.1 percentage points per annum over the decade 1996-2006 would contribute about 150 billion dollars to the deficit in 2006 and 700 billion dollars to the national debt by then (Advisory Commission to Study the Consumer Price Index 1996:...
been made over the years, there remain unsettled technical and theoretical issues that are extremely complex and difficult to deal with in the field of ‘measurement economics’. As one of the most distinguished experts of price index analysis recognises, ‘[d]ifficult technical issues inevitably arise in constructing price indexes, and such issues are often resolved by appealing to professional conventions’ (Pollak 1998: 74). In the case of the consumer price index (CPI), the most frequently used index to assess the movement in aggregate prices over a given span of time, recent work by the US Advisory Commission headed by Michael J. Boskin concluded that this index overstates the change in the cost of living by roughly one percentage point per year (Advisory Commission to Study the Consumer Price Index 1996). Researchers in the United States and elsewhere have come up with similar point estimates, with a range of plausible values of 0.5 to 1.5 percentage points per annum (see e.g. Cunningham (1996), Moulton (1996), Shapiro and Wilcox (1996), Greenlees (1997), Bureau of Labor Statistics (1997), Hoffmann (1998), and the huge literature cited therein). 'Hence, the very first point the CPI Commission made in its report was that inflation is inherently difficult to measure in a complex dynamic market economy’ (Boskin et al. 1998: 5). Further, while the CPI is the best measure currently available to gauge price movements, ‘it is not a cost-of-living index and it suffers from a variety of conceptual and practical problems’ (p. 23). In a nutshell, as Deaton has it, the CPI is ‘a concept that is hard to define and harder to measure’ (Deaton 1998: 38). It is precisely at the conceptual level that the present investigation

---

iii). In 2008, ‘the cumulative additional national debt from overindexing the budget would amount to more than $1 trillion’ (Boskin et al. 1998: 3).

2 With a total budget of 25,000 dollars, the Boskin Commission has probably written the most influential measurement paper of the century in terms of its impact. As Diewert states, ‘[e]very statistical agency in the world is reevaluating its price measurement techniques as a direct result of [the Advisory Commission's] report and the widespread publicity it has received’ (Diewert 1998: 56). Yet, as a corollary of the ‘Boskin Report’, a growing number of professional economists are becoming unsatisfied with the research strategy of measuring inflation via CPI computation. Nordhaus goes as far as to maintain that ‘[t]he CPI is so complex an organism that, like a Star Wars computer code, few can understand its exact functioning’ (Nordhaus 1998: 67).
is situated, after an initial survey of the technical issues raised by the measurement of aggregate prices\(^3\).

The second goal of this work, in fact, is to show that the recorded technical difficulties in measuring inflation by price index analysis stem ultimately from a yet imperfect macro-theoretic appraisal of the working of modern monetary economies of production. To quote from *The New Palgrave* dictionary article on inflation, by a prominent scholar of mainstream analysis, ‘[m]acroeconomics in general, and the theory of inflation in particular, is in a fluid state’ (Parkin 1987: 836). Based on the neoclassical analysis of money and inflation, the very concept of the price level seems indeed problematic. At the most fundamental level, it can be shown that any aggregate price index can logically be used to assess neither the purchasing power of

---

\(^3\) Issues of price measurement may be especially important for the European countries of the Economic and Monetary Union (EMU) area, and for their comparative performance according to the convergence criteria enshrined in the Maastricht Treaty (1992). This is why a Harmonised Index of Consumer Prices (HICP) has been developed over the last five years by Eurostat (1997). HICP was first used to monitor performance against the convergence criterion for price stability, and from January 1999 it is used by the European Central Bank (ECB) as the target measure of inflation for the EMU area as a whole. Indeed, within a big and structurally heterogeneous single currency area such as 'Euroland', with a unique, one-size-fits-all monetary policy across all participating countries, a single, consistently measured, aggregate price level is necessary for policy making and to gauge the area-wide economic developments. Yet, as the ECB’s President has been repeating since his appointment, the fact that the single monetary policy adopts a euro area-wide perspective means that it will not react to specific national shocks (Duisenberg 1998: 4). Clearly, this amounts to saying that the single monetary policy -- and thus the single, one-size-fits-all, short-term interest rate within the euro area, which is a necessary corollary of the European single currency -- will not in the event prove to be appropriate to the domestic needs of each of the euro-member ('in') countries. No-one doubts that such a risk exists. It may result from cyclical divergence within the euro area, with some participating countries needing to stimulate domestic demand while others are already operating close to capacity (see D. White (1999)). It may arise from differences in budgetary positions even though these are to be constrained through the Stability and Growth Pact (see Rossi and Dafflon (1999) on the loose definition of the latter). Or it may result from economic shocks of some sort that have a bigger impact on some countries than on others (e.g. the rise in oil prices in the 1970s, or the German reunification in the early 1990s). Eddie George, the Governor of the Bank of England, has recently been explicit on this point: ‘the risk of divergent monetary policy needs within the euro area is real. And if there were a material divergence of monetary policy needs, that could lead to serious tensions, because alternative adjustment mechanisms, such as labour migration or fiscal redistribution, that exist within individual countries, and which help to alleviate familiar regional disparities when they arise at the national level, are simply not well-developed at the pan-European level’ (George 1998: 2). A recent and interesting inquiry among the UK’s top academic economists, published in the April 17th issue of *The Economist* (1999), shows that these issues are a matter of concern for both the anti- and the pro-euro camp, although the two fronts attach different importance to them. These problems will be analysed in a later work.
money nor its variations over time. The principal argument is a compelling one, as we shall attempt to explain in detail. According to traditional analysis, monetarist as well as Keynesian, inflation is a situation in which too much money chases too few goods. This amounts to saying that any inflationary state is characterised by an excessive money stock compared to the total stock of saleable output. So far, so good. Yet, how is this inflationary gap (to be) measured? Reference to the general price level and its variation over time would be a logically correct one if, and only if, the two masses of money, on one side, and total output, on the other side, were autonomous and (at least in part) independent from one another. However, this cannot be so, neither in theory nor in practice. Besides the fact that the economic measure of output as a whole cannot be established independently of money (owing to the physical heterogeneity of the thousands of goods and services produced in any period), the aggregate level of prices is unable to measure objectively the two things on which it is based, namely money and output. Indeed, in the maintained method of economic analysis the measure of total output sold is determined via the aggregate price level, and the latter is arrived at by the confrontation on the market place between this very same output and the money stock. The argument is circular, as we shall see in the central part of the present work. Elaborating this point, it can then be shown that the theoretical problems of inflation analysis by means of price index computation may be ascribed to the underlying, and ill-grounded, neoclassical theory of money. As Lord Desai argues, one may confidently claim today that, in spite of two hundred years of monetary economics, there is still a manifest ‘lack of a theory for a monetary economy’ (Desai 1996: 1). The need therefore arises to go so far as to reconsider the

---

4 We shall examine the efficient causes of this phenomenon in due course.

5 This petitio principii notwithstanding, one must not overlook the fact that, by definition, each price level represents an equilibrium situation between money and output as a whole. So much so that by comparing the level of aggregate prices over time one can determine no inflationary disequilibrium at all.
nature of money, in order to understand inflationary pressures and, hopefully, to manage them accordingly.

The third, ambitious goal of this thesis is precisely to gain an understanding of inflationary disequilibria, grounding it on a sound theory of modern money. Assuredly, this is a challenging target. It might give rise to a number of fundamental questions which need further exploration, beyond the scope of this work. But it might also be far-reaching, inasmuch as it might provide us with a few macro-theoretic elements in order to elaborate an anti-inflation policy that can be successful. One of the most worrying conclusions arrived at by a number of inflation analysis surveys points indeed to the hazardous results of contemporary macroeconomic policies aimed at controlling inflation (Hudson 1982: 55). Despite the low levels of measured price increases recorded in recent years, it can in fact hardly be denied that inflation, at present, still is a pathology of our money-using economic systems. The very practice of indexation is in itself the sign of a disorderly working of our monetary economies of production. It is the purpose of this thesis to give a fresh look at this disorder, that is, inflation, from a relatively new macro-theoretical vantage point that goes so far as to reconsider money and banking principles. As stated by Cencini and Baranzini in their "Introduction" to *Inflation and Unemployment*, '[t]he concept of excess demand and the idea that inflation is provoked by an anomalous increase in the money supply have not been followed to their extreme implications and there is still a lot of confusion about the role played by banks in the process of money creation' (Cencini and Baranzini 1996: 1). It is along these lines that our analysis has been conceived in an attempt to shed some light on a yet obscure topic, in order to forge ahead a research programme capable of achieving monetary order. To be sure, much more work needs to be done in specifying the macro-theoretical elements put forward here,

---

6 As noted by De Vroey, '[b]efore examining the relationship of money to inflation, it is useful to dwell on the notion of money itself' (De Vroey 1984: 382).
before our story can be made fully persuasive. But to the best of our knowledge this is
the only attempt to integrate the economics of inflation within a monetary
investigation that goes to the roots of the phenomenon by critically assessing the
canonical research strategy of measuring, and seeking to control, inflation by means
of price index analysis.

Such an approach quite naturally requires this thesis to be divided into three parts
of equal importance. Following an order of increasing difficulty, this analytical
structure is intended to isolate the crucial features of contemporary inflation analysis,
in order to take the reader safely into a fruitful investigation of the political economy
of inflation.

The first part is devoted to the methodological issues in the measurement of
inflation, where emphasis is laid on both the perceived and unperceived weaknesses of
the conventional analysis of price level changes. The problems already start with the
need to define the object of inquiry, i.e. inflation. As soon as the latter is defined as
an ongoing increase in a somehow aggregate price level, the door is indeed opened for
the introduction of the multifaceted price index analysis and the technical-theoretical
conundrum these indices inevitably raise. Whilst the age-long methodological debate
in traditional analysis has focused on a number of technical issues, some important
conceptual problems have probably been unnoticed so far and are still waiting to be
brought forward in this framework. In particular, a principal point that seems to be
either misunderstood or neglected altogether in the ongoing debate on price
measurement issues concerns the heuristic status of a microeconomic investigation into
what fundamentally is a macroeconomic pathology, that is, inflation. To put it briefly,
the problems revolve around the paradigmatic approach to macroeconomics, linked to
the search for microfoundations. By assuming that the functioning, as well as the
malfunctioning, of the economic system as a whole can be apprehended, and perhaps
cured, by modelling the representative agents' forms of behaviour and then
aggregating the models' results over the entire set of economic units, mainstream
investigation has been led into a dead end⁷. This is particularly evident in inflation analysis, where it is traditionally assumed that (the loss in) the purchasing power of money can be assessed with a micro-statistical apparatus grounded on the aggregation of ‘elementary transactions’ adequately fitted into a quantity-theoretical ‘equation of exchanges’.

Although this work does not explicitly focus on the history of economic thought, the seeds of the main conceptual problems of contemporary inflation theory, and policy, can be found in the fundamentals of the quantity theory of money, which is indeed a quantity-of-money theory of the price level. Part II is therefore aimed at a critical appraisal of this theory, in both its ‘old’ and ‘new’ portrayal. The idea of the existence of a price level, as well as that of the possibility to determine its magnitude by applying index number theory, were in fact laid down by the founding fathers of twentieth-century quantity theory, among whom Irving Fisher stands out for his influential book on *The Purchasing Power of Money* (1911/1931) and his related work on *The Making of Index Numbers* (1922). Hence, as we shall argue at length, the way out of the dead end in current inflation analysis seems to require a renewed critical appraisal of the quantity theory. In particular, one is led to ask if money and output really are two distinct, and (at least partly) autonomous, things, and whether money can be depicted as a veil -- as suggested, for instance, by Professor Pigou (1949). Ultimately, the purpose of such an in-depth analysis of money is to provide some new insights into the hectic debate opposing the advocates of the exogeneity-of-money paradigm to adherents to the endogenous money approach. In this connection, the still widely shared belief that money owes its existence to the definition by the State of the unit in which payments are made (see e.g. Wray (1998b)) does not go far enough to be able to distinguish analytically the logical nature of money from either its historical or jurisdictional origins (see Realfonzo (1998) for an excellent discussion of these

---

⁷ See Ormerod (1994: Part I) for a methodological critique of the present state of economics.
issues with respect to the 1900-1940 money and banking debate\(^8\). A reconsideration of the means-of-payment function in adherence to a logical analysis of money is thus attempted in the central part of the present work, since this analytical step may be instrumental to elaborate a realistic, useful, and policy-oriented theory of inflation.

The analysis carried out in the third, and last, part of this thesis should then be considered as a modest attempt to contribute to the construction of a modern paradigm for inflation analysis that fully complies with the logical nature of money as well as with the factual working of a monetary production economy. Based on the line of thought developed by Bernard Schmitt and his School since the early 1950s, such an investigation is centred on the banks’ role in the monetisation of the national economy. In particular, a reconsideration of the essence of bank credit and its link to inflationary pressures may be pretty useful to re-appraise modern banking from a macroeconomic point of view. Starting from the intervention of the banking sector into what Davidson (1988) dubs the income-generating finance process, attention is drawn on the alleged monetisation by banks of agents’ conflicting claims on available output. Known as the conflict inflation approach, this theory puts emphasis on the supposed behavioural causes of inflation, whose origin is ultimately ascribed to an ongoing struggle for inconsistent income shares between the different functional categories of economic agents (namely, capitalists and labourers, who may be stratified according to their relative labour skills). Now, since the struggle for the distribution of national income cannot really modify the amount of the latter, it is a matter of fact that the money-output relationship is never influenced by the former. So, the next step is to ascertain whether or not the behaviour of the banking sector can

\(^8\) See also Bouvet (1996: 454) for a short discussion of the necessary distinction between the legal and the economic foundations of money and money’s worth within the endogenous money approach. A critical appraisal of recent endogenous-money literature can be found in Rossi (1999). This topic will not be addressed here, since the objective of this dissertation is to concentrate on a fundamental critique of traditional inflation analysis, grounded as it is on the concept of the price level and on the measurement of changes in the latter.
really generate inflationary pressures within the economic system. At this stage, care ought to be taken to assess the influence, both in space and in time, on the economy as a whole of the banks' decision to grant credit to the public. Provided a distinction is made between the micro- and the macro-level of the economy, it can be noticed that bank credit and overdrafts are two forms of advances, and as such they can never be much troublesome for the relationship between money and output. In other words, at the macro-level monetary equilibrium cannot be definitively altered by what microeconomic theories of the banks' behaviour consider to be an excess of credit creation. A further analytical step is required to grasp the malignant nature of inflation and its irremediable alteration of the relationship between money and output, as depicted by the loss in the purchasing power of money. Referring to the analysis worked out by Schmitt (1984, 1996b) and Cencini (1995, 1996), and particularly to their conception of 'empty money', it can be shown that inflation is a pathology that arises within the production process, as opposed to the circulation of already produced goods. From this vantage point, the origin of the loss in the purchasing power of money is to be found in the process of capital accumulation as recorded by the banks' bookkeeping. It is thus on the bookkeeping system of banks' transactions that research must be focused in order to see that the cause of inflation is 'structural' rather than 'behavioural'. As Cencini has pointed out, 'the relationship between money and output can be pathologically modified by a simple accounting mechanism that does not pay sufficient attention to the banking nature of money and to its functional link with production and circulation' (Cencini 1995: 70). It is therefore towards the elaboration of a monetary structure of payment systems fully complying with the fundamental requirements of bank money that our efforts have to aim. The research presented here is only a tentative step in this direction. Other, and more important, contributions are to be made if an adequate theory of a monetary economy is to come into existence, to manage inflationary pressures in such a way that monetary order can eventually be achieved. We hope nevertheless to have shown, as Lord Desai indicates in the
conclusion of his *Federico Caffè Memorial Lectures*, that 'it is possible to change old habits of thought and rethink the basic issues of monetary economics', to plant 'a seed that may bear fruit at some date in the not too distant future' (Desai 1996: 21).
PART I

METHODOLOGICAL ISSUES IN THE MEASUREMENT OF INFLATION
CHAPTER 1
THE METHODOLOGICAL DEBATE IN TRADITIONAL ANALYSIS

SOME NOTES ON THE CONVENTIONAL MEASURES OF INFLATION

Defining inflation

The problem of measuring inflation accurately starts with the attempt to define one of the distinguishing -- though not essential -- features which have been hampering the development of money-using economies from the antiquity through to modern times. As a number of surveys of inflation theory show, neither a satisfactory nor an exact definition of inflation exists as yet in economic literature, despite the age-long theoretical debate on the purchasing power of money which followed the publication of Ricardo's *The High Price of Bullion* in 1810. From early on, much effort about analysing inflationary disequilibria has indeed focused on the most evident factual outcome of the whole process, giving rise to a widely accepted pragmatic definition. Thus, in its most basic terms, 'inflation is a process of continuously rising prices, or equivalently, of a continuously falling value of money' (Laidler and Parkin 1975: 741).

Though commonly used as an undiscussed starting point for policy implementation, the symptom-based definition of inflation does not tell us anything about either the causes or the effects of the upward movement in prices. Following a pure empiricist perception, it simply underlines the fact that inflation manifests itself through a sustained increase in the general level of prices, that is, price increases have to be irreversible and concern a somehow weighted average of the prices of all goods and services sold in the marketplace, as we shall see later.

---

Yet, for those unwilling to accept the axiomatic equivalence between inflation and an ongoing rise in aggregate prices, the abundant literature offers more specialised definitions which bring out some particular characteristics of the phenomenon. In this framework, Bronfenbrenner and Holzman distinguish four alternative types of definitions:

1. Inflation is a condition of generalized excess demand, in which 'too much money chases too few goods.'
2. Inflation is a rise of the money stock or money income, either total or per capita.
3. Inflation is a rise in price levels with additional characteristics or conditions: it is incompletely anticipated; it leads (via cost increases) to further rises; it does not increase employment and real output; it is faster than some 'safe' rate; it arises 'from the side of money'; it is measured by prices net of indirect taxes and subsidies; and/or it is irreversible.
4. Inflation is a fall in the external value of money as measured by foreign exchange rates, by the price of gold, or indicated by excess demand for gold or foreign exchange at official rates (Bronfenbrenner and Holzman 1963: 599).

The first two definitions put forward a causal explanation of inflationary disequilibria, echoing the famous monetarist proposition that *inflation is always and everywhere a monetary phenomenon* in the sense that it is and can be produced only by a more rapid increase in the quantity of money than in output* (Friedman 1987: 17)*. According to both interpretations, the direction of causation goes from money supply growth to price increases ($\Delta M \Rightarrow \Delta P$), as was already claimed by the Bullionists at the time of the controversy between the Banking and the Currency Schools. In the first case inflation is traced back to demand-side factors, since the observed rise in prices is deemed to originate from excess aggregate demand in consumption, investment or government spending. In the second case a misleading behaviour of the monetary authorities is implicitly assumed to explain why an

2 Unless otherwise indicated, italics are from the original quotation.
excessive growth of the money supply occurs in modern economies. As Friedman put it, ‘[m]onetary authorities have more frequently than not taken conditions in the credit market -- rates of interest, availability of loans, and so on -- as criteria of policy and have paid little or no attention to the quantity of money per se. The emphasis on credit as opposed to the quantity of money accounts [...] for many of the post-World War II inflations’ (pp. 17-18).

The third definition quoted above is an elaborate version of the simple one laid down by Laidler and Parkin (1975). It suggests the existence of a kind of ‘natural’ rate of price increases similar to the hypothesis introduced by Friedman (1968: 8) of an underlying natural rate of unemployment\(^4\). It also considers inflation as resulting mainly from a (real sector) shift in the aggregate supply function due to production costs increases, thus stressing the alleged cost-push causal explanation (as traditionally opposed to demand-pull inflation)\(^5\). It finally introduces the analytical difference between anticipated and unanticipated inflation, which -- according to Laidler and Parkin (1975), Frisch (1983: 12), Parkin (1987: 833-835), and McCallum (1990: 964), among many others -- is the key distinctive feature of contemporary inflation theory.

The last definition recorded in the survey paper by Bronfenbrenner and Holzman (1963) concentrates instead on the international effects of inflation, which arise in individual countries that are open economies. Yet, as recognised by Parkin,

\(^4\) Much discussion about the relationship between inflation and unemployment has given rise to the familiar acronym NAIRU (standing for Non-Accelerating Inflation Rate of Unemployment), which is but a modern reformulation of the basic idea supporting the (long-run) Phillips curve (Phillips 1958). As neatly summarised by Gordon, '[i]f at any given time there exists a unique NAIRU, then the Phillips curve tradeoff is vertical at that unemployment rate' (Gordon 1996: 1).

\(^5\) It has been suggested by Frisch that the dichotomy separating cost-push and demand-pull inflation has lost its meaning in recent literature on inflation, because it is actually impossible ‘to identify empirically the two types of inflation’ (Frisch 1983: 12). Laidler and Parkin, following the same line of thought, go deeper and argue forcefully that '[s]ince inflation is a phenomenon affecting the whole economy, [...] the cost-push/demand-pull distinction [is] analytically unhelpful as a device for classifying those developments in inflation theory that are grounded in macro-economics' (Laidler and Parkin 1975: 742). This is a point to which we shall return later.
'Understanding the international generation and transmission of inflation in a flexible exchange rate world, such as that which had emerged by the mid-1970s, is still far from settled' (Parkin 1987: 835). In this work we shall not attempt to deal with this problem. The focus will be on the domestic aspects of inflation only.

Along with these 'empirical' definitions of inflation, it is possible to establish a threefold classification according to which criterion is used for that purpose (Table 1.1).

<table>
<thead>
<tr>
<th>Table 1.1 A typology of inflationary phenomena</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criterion for classification</td>
</tr>
<tr>
<td>-------------------------------</td>
</tr>
<tr>
<td>I. Causes of inflation</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>II. Expectations of inflation</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>III. Rate of inflation</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

*Source*: adapted from Frisch (1983: 11).

Each of these criteria has given rise to an immense literature, which is extensively dealt with, for instance, by Bronfenbrenner and Holzman (1963), Laidler and Parkin (1975), Frisch (1977, 1983), Hudson (1982), and Parkin (1987). Their original common ground, as formulated at the outset of this section, can easily be traced back to the traditional (neoclassical) dichotomy, according to which the notion of price level is established by relating the stock of money with the stock of output (Figure 1.1).
Figure 1.1 The concept of price level

In Figure 1.1 we put on a virtual ‘numerical’ scale the monetary mass, on one side, and the mass of domestic output (GDP), on the other side. The aggregate price level is arrived at by relating the two masses. A short numerical example may be useful here. Consider period 1 (for instance, December 1999), where 2,000 macro-units of money ($M$) exist alongside of 100 macro-units of output ($Q$) in any country whatsoever ($C_w$): the price level ($P = M/Q$) is then conventionally established at number ($P_1 = 20$). Now, in order to know whether there is inflation or not and, in the affirmative, to measure the inflationary pressure in the period under examination, standard monetary analysis requires the comparison with some period taken as a reference. Suppose that period 0 (say, December 1998) is the base-period, and assume that price level $P_0$ is equal to 18 (as expressed by the same convention as above). In its simplest form, dominant economic theory concludes that in the current period the observed national economy suffers from an inflationary gap whose measure ($\pi$) is extrapolated by the ratio of the two price levels, as in equation (1.1):

$$\left(\frac{P_1}{P_0} - 1\right) \cdot 100 = \pi.$$  \hspace{1cm} (1.1)

'The inflation rate, of course, is simply the rate of change of the aggregate price level' (Gordon 1992: 1). If the aggregate price ratio (i.e. $P_1/P_0$) is greater (smaller) than unity, we are told that in period 1 there has been an inflationary (deflationary) gap whose magnitude ($\pi$) is customarily expressed in percentage points, as in equation
So, in our example, a ratio of $10\%$ would mean that country $C_w$ witnesses an inflation of 11.1% over a twelve-month period, as calculated in December 1999.

Now, both the schematic representation in Figure 1.1 and equation (1.1) lead us directly to the problem of measuring the actual price level which ought to reveal the presence and the magnitude of any inflationary pressure. Let us begin with the canonical technical approach, which proves to be much more than a mathematical-statistical conundrum. We will address the related conceptual issues in the second part of this work.

**Introducing price index analysis**

The thousands of commodities produced in dynamic open economies such as those of OECD member countries make any price level measurement a formidable organisational and technical quandary. 'While it is conceptually easy to survey the prices of individual commodities at any given time, using these to produce a measure appropriate for monetary policy is far from straightforward. Gauging movements in aggregate prices is neither theoretically nor practically easy' (Cecchetti 1996: 1). Moreover, as Greenspan (1998: 3) has argued so convincingly, the notion of general price level -- and what one means by its change -- is never unambiguously defined. As a matter of fact, each household buys a different basket of goods and services over any given period of time, and most firms produce various sorts of heterogeneous commodities. In the field of 'measurement economics', as Triplett (1975: 19) calls it, '[i]ndex numbers are used to reduce and summarize this overwhelming abundance of microeconomic information' (Diewert 1987: 767).

Now, it is well known that the classical definition of index numbers can be traced back to Edgeworth, who proposed 'to define an index-number as a number adapted by its variations to indicate the increase or decrease of a magnitude not susceptible of accurate measurement' (Edgeworth 1925: 379). Applied to the study of inflation, the
clue of standard analysis as well as the origin of its methodological problems lie in this essential property of any price level. In particular, the crucial characteristic of index number analysis is that the number arrived at is only an index, not a measure, of the non-observable magnitude (Edgeworth 1925: 382; Allen 1975: 2). The calculation of the (general) price level is thus a mathematical-statistical process which can give only a roughly estimated signal about the movement in prices occurring between two points in time. This severe limitation of traditional investigation has already raised many doubts about its validity in measuring inflation accurately. For instance, the 1975 Report of the Inflation Accounting Committee is entirely against such a methodology. It maintains in fact that changes in the aggregate price level are actually unquantifiable and that, therefore, the compilation of a general index of price changes is of little practical use (Sandilands Committee 1975: 9). Moreover, as the joint winner of the 1975 Nobel prize for economics puts it in a classic paper on "Measurement Without Theory", in this framework the choices of ‘what measures to define and compute [for estimating price level changes] are made with a minimum of assistance from theoretical conceptions or hypotheses regarding the nature of the economic process by which the variables studied are generated’ (Koopmans 1947: 161).

Starting from the observed serious lack of satisfactory index numbers of the aggregate price level, Book II of Keynes's Treatise on Money deals indeed with the problem of measuring the value of money from a theoretical point of view. Keynes firmly believes that

[a]n Index-Number of the Purchasing Power of Money should include, directly or indirectly, once and once only, all the items which enter into final consumption (as distinct from an intermediate productive process) weighted in proportion to the amount of their money-income which the consuming public devote to them (Keynes 1930: 57).
In his analytical attempt to distinguish among a series of secondary price levels, Keynes brought to the fore the still fundamental problems of (a) how to measure the value of money in any national economy and (b) how to compare it between two distinct points in time or according to the state of the economy.

In the first place, we do not mean by Purchasing Power the command of money over quantities of utility. If two men both spend their incomes on bread and both pay the same price for it, the purchasing power of money is not greater to the one than to the other because the former is hungrier or poorer than the latter. [...] In short, comparisons of Purchasing Power mean comparisons of the command of money over two collections of commodities which are in some sense ‘equivalent’ to one another, and not over quantities of utility. The problem, therefore, is to find the criterion of ‘equivalence’ for this purpose (Keynes 1930: 96-97).

Foreshadowing most of present conceptions of cost-of-living indices -- which are based, as we shall see, on the theory of consumer demand and the implicitly related welfare approach6 (see e.g. Braithwait (1980), Pollak (1989), and the final report of the Advisory Commission to Study the Consumer Price Index (1996)) --, in his Treatise Keynes asserts that ‘comparisons of the purchasing power of money are the same thing as comparisons of the amounts of money-incomes of similar persons’ (Keynes 1930: 97). This framework of inquiry is basically of a microeconomic nature. Claiming that two sets of commodities are to be considered as ‘equivalent’ if the money-incomes spent for their purchase are the same for ‘two persons of equal sensitiveness’ (p. 97), it does not tell us anything on how to make such comparisons for the national economy as a whole.

6 ‘How much would it cost in today’s prices to make the consumer just as well off as he was yesterday? This question cannot be answered without resorting to an arbitrary intertemporal weighting of utilities’ (Fisher and Shell 1972: 1). To put the point sharply, Whittington -- referring to the work of Amartya Sen -- observes that ‘utility is strictly a subjective concept, and there is no objective method of comparing the pleasure, or “standard of living”, which different individuals derive from consuming even identical bundles of goods’ (Whittington 1983: 69). See also Keynes (1909/1983: 54).
We are in fact confronted with a double-sided approach to price level measurement, as neatly recognised by Allen (1975: 5-6), which might also explain in part the still hectic debate over the best measure of price movements. On the one hand, many academic researchers aim at a broad objective -- measuring the general price level or equivalently, so the argument goes, the purchasing power of money --, without specifying the reference sample of consumers or the group of producers considered in this 'stochastic approach'. On the other hand, the majority of empirical analysts follow an ‘aggregative approach’ whose objective is usually to study the net income or standard of living of a selected group of individuals (e.g. unskilled workers, elderly or rent-earning consumers).

Now, in this framework it need not be emphasised that ‘[a]s each household has a slightly different pattern of expenditure, price changes will have varying effects and households will experience different rates of inflation’ (House of Commons 1988: 2). A voluminous literature has dealt with this problem, too large to be surveyed by a single researcher. In practice, if not in theory, a price index number is a somehow extrapolated average, a synthetic figure derived from a great variety of price movements. As we are going to see, there are countless ways of computing an average. There are also ‘many different aggregate price measures, any of which could be used to measure the change in prices over time’ (Beaton and Fisher 1995: 5). The potential for arithmetical discrepancies between these different measures of inflation is indeed very high. How is it possible to choose among them the relevant price index for (monetary) policy implementation? Is there any ‘exact’ price index which can be used as an operational instrument to assess ‘true’ inflationary disequilibria from a macroeconomic point of view? Let us try to clarify the terms of the problem by referring to the technical path which has traditionally been followed to deal with it. In a further section we shall examine the usually associated methodological issues.
Selecting the appropriate formula

The calculation of the general price level and its related technical difficulties form part of the multifaceted index number problem (Samuelson and Swamy 1974; Allen 1975; Diewert 1976, 1987). Its practical outcome is a series of different formulae put forward by various authors in an attempt to evaluate the change in prices occurring between two situations (usually defined as two distinct periods, i.e. the base- or reference-period and the current- or end-period).

As is well known, the index number formulae most commonly used by contemporary statistical agencies to record the level of prices and the variations in that level are obtained from the original ones laid down respectively by Laspeyres (1871) and Paasche (1874). While attempting to improve the mathematical estimation of the average movement in prices in the national economy, Laspeyres was already aware 'that with these figures one cannot appraise the absolute loss in the purchasing power of money' (Laspeyres 1871: 302, our translation). Since the state of price statistics was not as reliable as it ought to be for that sort of extrapolation, Laspeyres maintained that price index analysis could only roughly estimate the decline in the purchasing power of money (p. 309).

Without dwelling on this technical issue from a historical perspective, the position of the Swiss Federal Statistical Office (SFSO) is particularly noteworthy for the case in point, inasmuch as the Swiss federal administration is going to introduce in year 2000 his sixth totally revised consumer price index since its first computation in 1922. Consider a quotation from the final report on the new conception of the Swiss CPI: 'A consumer price index may be defined and calculated in several ways. There is neither an ideal nor a unique index which is correct. Different conceptions are possible according to the use one is looking for' (Office fédéral de la statistique 1993:

---

7 The original (German) quotation sounds more peremptory: 'Dass mit diesen Zahlen über die absolute Grösse der Geldentwerthung nicht entschieden werden soll' (Laspeyres 1871: 302).
18, our translation). Which functional form for a price index should then be adopted, if we are looking for an accurate measurement of inflation from a macroeconomic standpoint? The SFSO (1993) report endorses the Laspeyres formula as the best solution for calculating national CPIs, more for the practical merits of its computation rather than for the theoretical qualities of the formula itself (p. 42). At a more sophisticated level, a large body of literature has grown up to work out the technical properties an index number formula should satisfy in order to justify its practical use. This 'axiomatic approach' brings to the fore a series of tests which seem to be reasonable and desirable for assessing the validity of economic index numbers (Diewert 1987: 768-769; Reserve Bank of New Zealand 1997: 22). 'The test approach to index number theory [...] looks at an index number formula from the viewpoint of its mathematical properties' (Diewert 1998: 48). By analogy with one-good-one-agent intertemporal equilibrium models, the criteria for finding well-behaved formulae say, for example, that

[i]f a single good's price doubles, the index should double; the index between any two dates will not be changed if the base period of the index is changed from one date to another; a dimensional change in the good (as from grams to pounds) should not change the index, nor should a dimensional change in money (as from pennies to dollars or dollars to pounds) (Samuelson and Swamy 1974: 566).

Now, it is a well-established fact that 'the Laspeyres price index does not represent an ideal solution, neither from the viewpoint of both the axiomatic and the economic approaches, nor from the viewpoint of its use for deflating national accounts' (Office fédéral de la statistique 1993: 42, our translation). Indeed, '[the

---

8 No translation can properly render the emphatic tone of the original (French) version: 'Un indice des prix à la consommation peut être conçu et calculé de nombreuses manières. Il n'y a pas d'indice qui soit idéal ou le seul à être juste. Différentes conceptions seraient parfaitement justifiables suivant l'utilisation prévue' (Office fédéral de la statistique 1993: 18).
9 'L'indice des prix de Laspeyres ne constitue une solution idéale ni du point de vue des approches axiomatic et économique, ni du point de vue de son utilisation comme déflateur pour la comptabilité nationale' (Office fédéral de la statistique 1993: 42).
Laspeyres index assumes no consumer substitution occurs in response to changes in relative prices, an assumption that is extreme, unrealistic and unnecessary’ (Boskin et al. 1998: 7). On purely mathematical grounds, several alternative formulae have been put forward to give a better estimate of changes in aggregate prices than the one laid down by Laspeyres, which can be written as equation (1.2):

\[
\frac{\sum_{i=1}^{n} p_i^1 q_i^0}{\sum_{i=1}^{n} p_i^0 q_i^0} \times 100 = P_L,
\]

where: \( i = \text{commodity } i \) (for \( i = 1, \ldots, n \))

\( p_i^1 \) = price of commodity \( i \) in (current-) period 1

\( p_i^0 \) = price of commodity \( i \) in (reference-) period 0

\( q_i^0 \) = quantity of commodity \( i \) purchased in (base-) period 0

\( P_L \) = Laspeyres price index.

Addressing the choice of the relevant quantity vector (the \( q_i \)'s) for price index analysis, Paasche argued that to reflect actual inflation this vector should capture the structure of expenditures in the current period (Paasche 1874: 171-173). The statistical approach suggested by Paasche associates therefore both price vectors with the quantities pertaining to the end-period, as in equation (1.3):

\[
\frac{\sum_{i=1}^{n} p_i^1 q_i^1}{\sum_{i=1}^{n} p_i^0 q_i^1} \times 100 = P_p,
\]

where: \( q_i^1 \) = quantity of commodity \( i \) purchased in (current-) period 1

\( P_p \) = Paasche price index.
Now, as a general rule, it is usually claimed that the Laspeyres index tends to overstate the upward movement in prices, while using the Paasche formula leads to underestimating price increases\(^{10}\) (see e.g. Triplet (1975: 22), Frisch (1983: 14-15), Gordon (1992: 7), Cunningham (1996: 13), Boskin et al. (1998: 6-7), and Diewert (1998: 48)). Further, according to one of the most prominent researchers of the US Bureau of Economic Analysis, 'the ratio of two different Paasche indexes has no standing in the theory of index numbers, and cannot be interpreted as an inflation measure' (Triplet 1980: 569). In particular, calculating the change in a Paasche price index between two points in time (say, period \(t+1\) in comparison with period \(t\)) has no clear meaning as a measure of inflation, because this ratio would have two distinct sets of weights (the \(q_i's\) in equation (1.3)), as in expression (1.3*)\(^{11}:\)

\[
\frac{P_{t+1}'}{P_t'} = \frac{\sum_{i=1}^{n} p_i'^{t+1} q_i'^{t+1}}{\sum_{i=1}^{n} p_i^0 q_i'} \frac{\sum_{i=1}^{n} p_i^0 q_i'}{\sum_{i=1}^{n} p_i q_i'}
\]

More recent theoretical work on index numbers has put forward different 'mixed-weight' formulae, in order to provide closer estimates of 'true' price movements in the aggregate (see especially Diewert (1987) for a survey). As postulated by Pigou

\(^{10}\) The result that the Laspeyres price index always exceeds the Paasche price index is critically dependent on the functional assumption that consumers' preferences are homothetic (Blow and Crawford 1999b: 11). 'In general, if homotheticity is violated, the Paasche price index may actually exceed the Laspeyres price index' (Anderson et al. 1997: 47). In practice, if not in theory, homotheticity implies unitary income elasticities of consumers' demand functions, i.e. that individual preferences for different goods do not depend on the households' level of welfare (Samuelson and Swamy 1974: 566; Deaton and Muellbauer 1980: 142-145; Whittington 1983: 68; Barnett 1987: 145-149).

\(^{11}\) By contrast, the economic interpretation of the rate of change of a Laspeyres price index is straightforward, for this ratio is defined with a single (base-period) quantity vector (i.e. \(q_i^0\), as in equation (1.2)). Thus, after simplification, the Laspeyres version of expression (1.3*) would be written as follows:

\[
\frac{P_{t+1}'}{P_t'} = \frac{\sum_{i=1}^{n} p_i'^{t+1} q_i'^{t+1}}{\sum_{i=1}^{n} p_i^0 q_i^0}
\]
Fisher (1922), and Bowley (1928: 217) in their contributions to the in-depth research activity on index numbers of the 1920s, a geometric average of $P_L$ and $P_p$ might represent a better approximation of the underlying aggregate price level, of which the Laspeyres and Paasche measures could be seen, respectively, as the upper and the lower estimated bounds. In its original form, what is now called the 'Fisher ideal price index' ($P_F$) is given by the square root of the product of $P_L$ and $P_p$ (equation 1.4):

$$\sqrt{P_L \cdot P_p} = P_F$$ (1.4)

Further derivation from this 'superlative index number formula' -- according to the terminology first proposed in a seminal paper by Diewert (1976: 136-137) -- led a number of mathematical statisticians to extrapolate price changes from a weighted geometric mean of the growth rates in prices\(^{12}\), by adopting the functional form advocated by Törnqvist (1936: 28) (equation 1.5)\(^{13}\):

$$\prod_{i=1}^{n} \left(\frac{P_i^1}{P_i^0}\right)^{s_i} \cdot 100 = P_T$$ (1.5)

where: $s_i = (0.5) \frac{P_i^0 q_i^0}{\sum_{i=1}^{n} P_i^0 q_i^0} + (0.5) \frac{P_i^1 q_i^1}{\sum_{i=1}^{n} P_i^1 q_i^1}$ represents the average expenditure share on good $i$ over the two periods.

In fact, neither $P_F$ nor $P_T$ (the Törnqvist price index) has ever been widely used in the economic field, mainly because these formulae have no direct interpretation in

\(^{12}\) 'A superlative index requires the same information on prices and quantities as a fixed weight index, but involves interpolating between the two periods rather than treating one of them as the "base" period' (Advisory Commission to Study the Consumer Price Index 1996: 23).

\(^{13}\) See Diewert (1987: 768). I am indebted to Marcello Corti, of the Swiss Federal Statistical Office, for pointing out a formal slip in an earlier version of this chapter.
practice (Allen 1975: 2; Frisch 1983: 15; Fortin 1990: 109-110). On the other hand, the continuous-time price index number due to Divisia (1925, 1926) -- which also inspired the establishment of $P_T$ as openly recognised by its proponent (Törnqvist 1936: 28, n. 2) -- is well known among econometricians for its many desirable properties (see e.g. Richter (1966: 749-753)). However, as testified by econometrics literature, the basic difficulty with Divisia indices lies in what has been termed the path independence problem (Hulten 1973: 1023-1024; Samuelson and Swamy 1974: 578-580; Diewert 1976: 124-129). Being obtained by a line integration over the path followed by the individual variables (in our case, prices and quantities entering the market basket over time interval $[0, t]$), the Divisia price index depends, as a general rule, on the trajectory over which the integration is taken. This means that 'a multiplicity of index values may be associated with any given point in the set of variables being indexed' (Hulten 1973: 1017), depending on the particular path of integration which has been selected. Yet, to obtain the necessary and sufficient conditions for path independence, the Divisia index requires a set of very restrictive assumptions which threaten its analytical validity (Hulten 1973: 1018-1019; Griliches 1990: 188). In particular, to avoid indeterminacy, one has to assume that the vector-valued function is linearly homogeneous, and that, for empirical analysis, the continuous-time (shifting) weights can be approximated by some kind of either discrete (two-period) averages or chain indexing procedures (Hulten 1987: 900).

All in all, without going further, we can conclude that the technical-statistical approach does not enable us to define any aggregate price measure as being better than the others. From a theoretical perspective, it is impossible to discriminate among $P_L$, $P_P$, $P_F$, $P_T$ and whatever other price index formula the large literature may provide. Many contemporary leading professionals (econometricians, statisticians, and political economists) are indeed well aware of this failure, but do not seem to worry too much. Diewert is a case in point. According to him, 'it does not matter very much which of these formulae we choose to use in applications: they will all give the same
answer to a reasonably high degree of approximation’ (Diewert 1987: 773). Yet, given the widespread use of price indices in current economic analysis -- from monetary policy decisions to assessing national economic performance (and related cross-country comparisons) and escalating government spending and taxes --, it can hardly be doubted that the exact measurement of inflation is a key problem which deserves a careful and thorough investigation, if we want to find an answer to the provocative assertion that ‘the opinion of representative house-wives would be preferable to the formulae of mathematical statistics’ (Edgeworth 1925: 380).

DEBATED METHODOLOGICAL ISSUES IN TRADITIONAL ANALYSIS

Since the publication of the so-called ‘Stigler Report’ (National Bureau of Economic Research 1961), which represents the most comprehensive review of American price statistics in the twentieth century, the official price indices compiled regularly by government statistical agencies have been subjected to a persistent barrage of methodological criticism (Blinder 1980: 539; Gordon 1990: 9). As a matter of fact, ‘[t]here has been increasing interest in recent years in the extent to which official consumer price indices may mismeasure the true rate of inflation’ (Blow and Crawford 1999b: vii). Although the vast empirical literature on this topic focuses mainly on three G7 countries (namely, the United States, Canada, and the United Kingdom), the same set of issues apply, at least in part, to other official price indices as well (e.g. the newly revised Swiss CPI, which may be considered in the present context as representative of a heterogeneous, though small, country). How to obtain

---

14 This also according to the Advisory Commission to Study the Consumer Price Index (1996), which ‘did not have the substantial resources that […] the so-called Stigler Commission […] had in 1961’ (p. 88).

15 It is a fact that a representative ‘Swiss consumer’ does not exist, despite the limited size of the Swiss economy. Consumption is closely related to the culture, traditions, lifestyles, and development of the different regions forming the Swiss landscape. The same contention may be made, at a higher level, with respect to the European Union, since there is no identifiable ‘European consumer’ as such. The recently established Harmonised Index of Consumer Prices for EU countries participating (or
information on who is buying what, where, when, why and how in an economy, and then to aggregate it into one or a few measures of price change raises a host of complex analytical and practical problems' (Boskin et al. 1998: 9).

With few exceptions, all aggregate price statistics are still constructed using fixed-weight Laspeyres indices, with reweighting every five to ten years (Diewert 1987: 773; Turvey et al. 1989: 38). Table 1.2 shows the frequency of weighting revisions within the G7 (in comparison with a small, yet heterogeneous, country such as Switzerland), but it should be recalled that in some cases the process of revising weights does not abide by a strict temporal rule (Italy is a typical example).

<table>
<thead>
<tr>
<th>Country</th>
<th>Index name</th>
<th>Frequency of weighting revisions</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States</td>
<td>Consumer Price Index (CPI)</td>
<td>10 years</td>
</tr>
<tr>
<td>Japan</td>
<td>Consumer Price Index (CPI)</td>
<td>5 years</td>
</tr>
<tr>
<td>Germany</td>
<td>Consumer Price Index (CPI)</td>
<td>5 years</td>
</tr>
<tr>
<td>Canada</td>
<td>Consumer Price Index (CPI)</td>
<td>4 years</td>
</tr>
<tr>
<td>Italy</td>
<td>Consumer Price Index (CPI)</td>
<td>3 to 4 years†</td>
</tr>
<tr>
<td>France</td>
<td>Consumer Price Index (CPI)</td>
<td>yearly</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>Retail Prices Index (RPI)</td>
<td>yearly</td>
</tr>
<tr>
<td>Switzerland</td>
<td>Consumer Price Index (CPI)</td>
<td>5 years</td>
</tr>
</tbody>
</table>

†The Italian weights are not revised following a temporal rule, but were revised in 1985, 1989, 1992 and 1996.

Sources: Office fédéral de la statistique (1993: 43); Advisory Commission to Study the Consumer Price Index (1996: 16); Cunningham (1996: 42, Table 6).

wishing to participate) in the final stage of European Monetary Union — which is ‘in the form of an index covering the euro area as a whole’ (European Monetary Institute 1997: 75) — is grounded in fact on a highly hypothetical consumption behaviour, based on a representative basket of consumers' goods averaged over many different, heterogeneous households. See Eurostat (1997) for both the methodological and the legal framework of HICP.
Such an approach to price level measurement is necessarily backward looking (Heymann and Leijonhufvud 1995: 99). Since aggregate price indices are compiled regularly but with a time lag and, moreover, with weights that fail to reflect truly contemporaneous consumption patterns\(^\text{16}\), they cannot be used to measure the movement in current prices. Further, although there already are many practical difficulties in gathering complete and accurate price data at the level of both representative items and elementary aggregates (Allen 1975: 39; Turvey et al. 1989: 53-83; Foss 1993: 277-278), a major methodological problem lies in the impossibility to obtain price information for the same bundle of commodities and services through time (Armknecht and Weyback 1989: 107). As pointed out by the US Price Statistics Review Committee headed by George J. Stigler,

> [t]he data used in computing the value of a price index are ordinarily derived almost entirely from a highly complex network of samples -- samples of goods and services, samples of localities in which prices are collected, samples of actual price reporters, and samples of points in time. It is therefore apparent that a value of an index depends upon the particular samples from which the basic data are obtained, and that different samples will lead to possibly different values of the index (National Bureau of Economic Research 1961: 39).

On the whole, this state of the art is not altered but strengthened by the fact that commonly used price indices contain many sorts of biases, that are nevertheless to be kept conceptually distinct from the more easily avertible computational errors (which in practice have been eradicated since the introduction of both high-performing computers and well-tried software). For a caricatural picture of this latter aspect, let us follow Gordon's (1981) evocative tale and imagine that someone pushes the wrong button on a computer used for CPI calculation. If the resulting estimation of 'headline' inflation rate (i.e. the twelve-month change in the overall CPI) is, say,

\(^\text{16}\) As Boskin et al. notice, 'the expenditure weights used in the Consumer Price Index are several years out of date even on the first day a revision is introduced. For example, [in the United States] the 1982-84 weights were implemented in the CPI program in 1987, and the 1998 revision will use weights from 1993-1995' (Boskin et al. 1998: 7).
12.1% instead of the ‘true’ rate of 11.1%, as recorded by a competing government price measure such as the ‘Gross Domestic Product deflator’, then millions of households and firms may undergo a number of extra charges, or may benefit from windfall gains, which would also have important budgetary consequences for the general government sector. To cite but an example taken from a recent US Congressional Budget Office Paper, ‘[i]f the CPI has an upward bias, some federal programs would overcompensate for the effect of price changes on living standards, and wealth would be transferred from younger and future generations to current recipients of indexed federal programs’ (quoted by the Advisory Commission to Study the Consumer Price Index (1996: 6)).

No-one doubts that the construction of a measure for aggregate price movements still is a politically sensitive area (Gordon 1992: 1; Silver and Ioannidis 1994: 555; Shapiro and Wilcox 1996: 3). The different distribution of income resulting from different price level estimates is perhaps the main obstacle to reach a broad political consensus on the extent to which money wages, social security expenditures and other disbursements (e.g. separate maintenance) ought to be escalated (by the annual percentage change in CPI) by law or by private contract. Any statistical extrapolation from official price series is likely to find both supporters and opponents, because, as we noted in the previous section, no single approach to index number construction yields the ‘optimal’ functional form (Allen 1975: 3; Diewert 1987: 767-768). In the wake of the recent debate on inflation measurement issues, especially in the United Kingdom and in the United States, this might be a good time to re-examine the problem of inflation from a relatively new macroeconomic perspective. Let us first

---

address the methodological problems of traditional investigation and begin with the most debated issues. We shall turn to the deeper analytical questions in Parts II and III.

The compositional issue of the representative market basket

Generally speaking, it is claimed that the CPI is the best measure currently available for measuring inflation (see e.g. the final report of the Advisory Commission to Study the Consumer Price Index (1996: ii), as well as Cunningham (1996: 9) and Steindel (1997)). The theoretical framework for CPI construction is provided by the theory of consumer demand, as developed for instance by Lancaster (1971, 1991), and Deaton and Muellbauer (1980). Demand and hence the determinants of prices are based on the canonical utility function approach, first translated in indifference maps on a bidimensional (Cartesian) system by Edgeworth. Similarly, CPI compilation is grounded on a basket of consumption items (goods and services) sold in the marketplace, whose quantities (i.e. the ‘weights’) are fixed according to the most recent nation-wide household expenditure survey\(^{18}\). To this approach, Keynes’s main (involuntary?) contribution was to argue, in *A Treatise on Money*, that

\[ \text{[s]ince the Purchasing Power of Money in a given context depends on the quantity of goods and services which a unit of money will purchase, it follows that it can be measured by the price of a composite commodity, made up of the} \]

\(^{18}\)In the United States, the main source of information used to determine the specific items which comprise the overall market basket is the Consumer Expenditure Survey (CES). ‘The Consumer Expenditure Survey provides a continuous and comprehensive flow of data on the buying habits of American consumers for use in a wide variety of economic research and analysis, and in support of revisions to the Consumer Price Index’ (Mason and Butler 1987: 22, n. 3). Since both the material and financial efforts of carrying out this review are very expensive, the federal government ‘is only willing to allocate funds for such a survey every decade’ (Gordon 1981: 116-117). Statistics Canada, the Canadian central statistical agency, conducts every four years a Family Expenditure Survey (FES), whose results provide the expenditure data for weighting the national CPI (Fortin 1990: 116). In the United Kingdom, ‘[t]he FES is an annual random cross section survey of around 7,000 households (this represents a response rate of around 70%). [It] records data on household structure, employment, income and the spending over the course of a two week diary period’ (Blow and Crawford 1999a: 16-17). See also Beaton and Fisher (1995: 6) and Cunningham (1996: 17).
various individual goods and services in proportions corresponding to their importance as objects of expenditure (Keynes 1930: 53).

Now, in principle, an 'ideal' aggregate price measure would logically include all economic transactions recorded over the chosen period (e.g. a calendar month). In practice, however, to make price index analysis manageable, government statistical agencies apply 'a simplified view of the marketplace and consumer behavior. This simplified view is reflected throughout the CPI approach' (Advisory Commission to Study the Consumer Price Index 1996: 11). Thus, by definition, the selected sample of goods and services included in the CPI's market basket covers but a part of currently produced output and varies from a list of about 300 representative items (such as in the Swiss case) to more than 70,000 commodities (as those subsumed in the American overall CPI).

For instance, an important category of expenditures often omitted from CPI calculation, because of the 'almost insurmountable difficulties regarding existing methodology' (Balk 1980: 68), concerns some goods and services of an exceptionally seasonal nature, such as fresh fruits and vegetables, Christmas trees or summer-rented flats. As testified by empirical work of a number of both professional statisticians and econometricians, '[f]or food items, the price dispersion within an RPI item in a particular month, as measured by the coefficient of variation, is typically 5-15 per

---

19 It may be noted in passing that, by definition, this approach (stemming, basically, from a microeconomic research programme) cannot integrate the submerged (black) economy in the statistical model it rests upon (see e.g. Siesto (1987) for an attempt to evaluate the submerged economy as a percentage of GDP). Failure to do so adds to the list of methodological drawbacks of price level measurement.

20 Ufficio federale di statistica (1993: 4). The British RPI is based upon 500-600 goods and services (see e.g. Craven and Gausden (1991: 29), and Beaton and Fisher (1995: 7)).

21 Actually, the price data for some 70,000 to 80,000 goods and services surveyed in the United States are used to form price indices for 207 item groups, which in turn are aggregated to form the overall CPI. See e.g. Advisory Commission to Study the Consumer Price Index (1996: 24-25), and Hulten (1997: 92), for more details.

22 'When making policy, central bankers would like to avoid responding to seasonal fluctuations in price data. While seasonality may be easy to understand in theory, it is extremely difficult to actually remove from most economic time-series' (Cecchetti 1996: 16). For a recent methodological analysis of the seasonal component of aggregate price series in Italy, see Cubadda and Sabbatini (1997).
cent and for some foods such as fresh vegetables and some meat may be 25 per cent’ (Carruthers et al. 1980: 20).

In a similar vein, the Bureau of Labor Statistics (BLS) -- the government statistical agency responsible for CPI calculation in the United States -- omits deliberately two categories of goods and services, namely ‘out of scope’ and ‘truncated’, from the overall consumer basket (Gale 1981b: 60). Out-of-scope items (e.g. commuting expenses and life insurance) are not included in the Consumer Expenditure Surveys, mainly because the BLS has not been able to determine how to measure them for price index analysis. Truncated goods data, instead, are collected in those surveys but do not enter the actual sample of representative items forming the CPI’s market basket23.

23 The category of ‘truncated goods’ is excluded from the consumer price index on account of its non-representativeness of households' consumption. It may comprise, for example, club subscriptions, primary and secondary private schooling, administrative charges for credit-card accounts, musical instruments, and some purchases of jewellery (Gale 1981b: 60; Turvey et al. 1989: 11-14; Ufficio federale di statistica 1993: 13). Its exact content varies, of course, according to the country considered. It has also been alleged that ‘nonmarket items that heavily affect the quality of life are generally beyond the current scope of such indexes’ (Boskin 1996: 23). What is particularly noteworthy here is the difficulty of including public goods and services in CPI measurement, for, as a general rule, there are very few observable market prices in the public sector. Another related issue concerns the depletion of non-renewable resources and other unmeasured items such as environmental services. In fact, in the framework of ‘green’ national accounting, ‘the value that consumers place on the current level of all nonmarket services provided by the environment, presents severe measurement difficulties’ (Hamilton and Atkinson 1996: 682). Yet, granting the importance of providing an all-inclusive socio-economic indicator for the actual level of human well-being as well as for sustainable development (see Pearce et al. (1996), and the environmental economics literature cited therein), the problem of ‘greening’ national accounting aggregates does not seem to be relevant to our investigation. Both environmental services (such as unpolluted air and water) and natural resources (for instance, oil and fuelwood availability) are not part of the economic field – unless they are monetised within a production process. To that extent, they logically do not have to be considered when enquiring into the (alteration of the) relationship between money and output (i.e. money’s purchasing power). A similar reflection also applies to unpaid work -- for instance, a housewife’s priceless contribution to the household’s welfare or, at a more trivial level, do-it-yourself activities like gardening, decorating, plumbing etc. (see Robinson (1956: 17) for a short list of important, but non-monetised, activities in modern societies). Although all these activities involve human effort (Jevons’s disutility of labour), the fact that they actually do not generate money income excludes them from the field of inflation measurement. We should however not lose sight of the fact that all environmental as well as domestic services ought to be monetised, in order to measure resource depletion and pollution emissions in social terms, on the one hand, and to value yet unpaid work, on the other hand. Luxton (1997) provides a good survey of the problems involved in measuring and valuing unpaid work, although it would carry us too far afield to explore them in any further detail here.
Yet, even accepting this framework of inquiry, the problem is, as already recognised in *A Treatise on Money* by Keynes himself, that ‘the composite commodities representative of the actual expenditure of money incomes are not stable in their constitution as between different places, times or groups’ (Keynes 1930: 95). By adopting the methodology of statistical sampling, the agencies responsible for official CPI computation are therefore constrained to accept, quite mechanically, an unrealistic set of restrictive (behavioural) assumptions, which are reflected in the sampling techniques for selecting (a) the relevant geographic areas, (b) the reporting outlets, (c) the time of data collection, and (d) the population surveyed. As recognised by the authors of the so-called Boskin Report, ‘one has to define the commodities and services the prices of which one wants to measure, how to measure them, how to collect data on them, over what span of time and at what interval, where and when to collect the data, and how to aggregate them into one or several

---

24 As we shall begin to see in the next chapter, the resulting discrepancies with respect to macroeconomic reality are more fundamental than ‘measurement biases’.

25 According to the ‘Boskin Report’, the BLS collects price data in 88 locations, called Primary Sampling Units (PSUs). In fact, only the price data of 44 geographic areas are aggregated for monthly CPI compilation (Advisory Commission to Study the Consumer Price Index 1996: 13, n. 10). In a similar mood, with the fifth (1993) CPI reform the Swiss Federal Statistical Office has halved the number of municipalities (‘communes’) where price data are collected. There are at present (2000) 24 geographic areas which should represent the whole Swiss economy for price index analysis. Office fédéral de la statistique (1993: 35-36, 79) and König (1995: 37-39) give more details on this specific point.

26 ‘Taking into account both the dispersion of price movements and the structures of distribution of particular goods, one needs as a general rule 2 to 6 price collection points per commune and for each type of goods’ (Office fédéral de la statistique 1993: 36, our translation). [‘Compte tenu de la dispersion des variations de prix et des structures des canaux de distribution de certains produits, il faut en règle générale 2 à 6 points d’observation des prix par commune et par type de produits’ (Office fédéral de la statistique 1993: 36).] For its part, the American CPI rests upon expenditure data collected from 22,000 outlets sampled according to the so-called Point-Of-Purchase Survey (POPS) (Advisory Commission to Study the Consumer Price Index 1996: 12-13).

27 In the Swiss case, price data are collected during the first week of the calendar month to which the monthly CPI refers, in order to be as timely as possible (Office fédéral de la statistique 1993: 40).

28 There are indeed as many possible CPI as there are different groups of consumers. In the United States, the two most commonly used measures are CPI-U and CPI-W. ‘The former is for all urban consumers, roughly 80% of the population; the latter is for urban wage [earners] and clerical workers, about 32% of the population’ (Advisory Commission to Study the Consumer Price Index 1996: 2, n. 2). This is an important point to which we shall return.
overall summary statistics. At each of these levels, various judgments and assumptions must be made to make practical headway' (Boskin et al. 1998: 6). These judgments and assumptions are epitomised in the concept of the fixed-weight market basket, which necessarily becomes less and less representative over time as consumers buy different bundles of commodities (Advisory Commission to Study the Consumer Price Index 1996: i). Indeed, the most striking example of the methodological problems inherent in the CPI approach probably relates to the failure to capture new sorts of goods in a timely manner. This built-in deficiency reduces the ability of price index analysis to cover the actual consumption basket and, as we shall see later, to measure inflation accurately.

Now, a further difficulty in defining the appropriate market basket for inflation analysis arises with the category of 'durable goods'. As one of the most distinguished scholars in 'measurement economics' puts it, this is 'the segment of the economy where official price data are most vulnerable to inaccuracy, owing to the heterogeneity and changing specifications of durable goods' (Gordon 1990: 9). For example, it is a well-known fact that, in developed market economies, by far the largest single weight in the overall CPI is given to the housing component (which is itself composed of several expenditure classes). Without going into the details of this category, it is sufficient here to mention the twin major controversial points in the treatment of housing expenditure data, namely the treatment of homeownership costs and the related mortgage interest payments. In this respect, as reported for instance by Gordon, the American government statistical agency before 1983 had been making

---

29 One cannot be more explicit than this: 'The United States became a motorized society in the 1920's and 1930's, when there was an enormous improvement in the performance of automobiles along with a decline in their price – but the automobile was not included in the CPI until 1940' (Gordon 1981: 130). As a more recent noteworthy example, in 1996 there were approximately forty million mobile phones in use in the United States, but the mobile phone was not included in the American CPI's representative basket at that time (Advisory Commission to Study the Consumer Price Index 1996: 39).

30 In the United States, Canada, the United Kingdom, and Switzerland, the housing component accounts for at least one-quarter of the overall CPI.
'the fatal error of treating the whole population as if it were in the predicament of a newlywed couple buying its first house' (Gordon 1981: 121). In effect, the BLS treated each homeowner as if he or she were buying a house outright, since it used price data of newly constructed home purchases for estimating the rate of change of the CPI's dwelling costs (Blinder 1980: 550-552; Gordon 1992: 40-41). In 1983, the housing component of the American CPI shifted to the so-called 'rental equivalence' approach, which simply assumes that homeownership costs move in proportion to market rents observed in the domestic private sector for similar properties. The same treatment had been applied in the United Kingdom until 1975. In this country the costs of homeownership in terms of shelter prices have never been directly included in the monthly-calculated RPI. 'Before 1975 owner occupiers' housing costs (other than rates and maintenance) had been represented in the RPI by estimates of the rents which properties might have been let at on the open market' (Evans 1989: 39). Yet, because of practical contingencies engendered by a marked slowdown in the British private-rented sector during the 1970s, since 1975 homeownership costs entering the monthly RPI have been measured by referring to annual mortgage interest payments, thus resorting to the well-known 'user cost' concept. There were in fact many methodological objections to the standard rental-equivalence formulation, centred around the fact that most single-family homes are not rented and, therefore, the data collected in open markets cannot be used to estimate virtual rents of single-family houses (Blinder 1980: 555; Gordon 1981: 126; Wynne and Sigalla 1996: 69). However, estimating housing expenditure data by adopting user cost formulae does not really improve the methodology of price index analysis. As far as mortgage

31 A number of OECD countries, such as Germany, Spain, the Netherlands, and New Zealand, rely upon similar properties' market rents for estimating the housing costs actually supported by owner-occupiers. See Evans (1989: 54, n. 6) for bibliographical references; Craven and Gausden (1991: 32-37), and Reserve Bank of New Zealand (1997: 7-10) for further details.

32 On measuring homeownership costs by user cost formulae, see Blinder (1980: 557-565) and the literature he refers to.
interest disbursements are concerned, the calculation of the official homeownership index is indeed questionable, because ‘the amount borrowed by a household for house purchase is largely determined not by house prices but by household income’ (Evans 1989: 41). Further, the user cost approach is too ambiguous to give a clear-cut shelter price proxy for the owned-accommodation index (Blinder 1980: 560-561; Evans 1989: 49-51; Fortin 1990: 125-127; Gordon 1992: 41). There are in fact many ways to measure the user cost of housing, because various components might be included in (or excluded from) the relevant equation and, moreover, each ingredient (particularly, the mortgage interest rate) turns out to be inherently volatile.

On the whole, although the rental equivalence approach might seem to fit the CPI’s housing component better than the user cost approach, there are many serious problems of both a methodological and an empirical nature in its implementation. These problems are not specific to the housing component of aggregate price indices. An analogous set of issues may indeed apply to most durable goods (e.g. cars, aircraft, computers, and other electrical appliances such as hi-fi equipment or microwave ovens -- to quote only the main categories where empirical work is being carried out for price measurement purposes), for the problem at stake reverts to the daunting, and multifaceted, stock-flow analysis. In short, in the jargon of economics, the widely accepted definition of durability implies that one has to decide if, and how, to distribute the value-in-use of a durable good over its lifespan. The locus classicus of all such analyses is ‘the fact that durable goods are consumed (and hence yield utility) only gradually over time’ (Blinder 1980: 549). Indeed, many influential economists maintain that we must price the service flow (say, over a year) of durable goods, instead of adopting their purchase price, when compiling price indices for

33 The British government is indeed well aware of these failures, and tries to overcome them by targeting the annual growth rate of an index (known as RPIX) which excludes -- attaches a zero weight to -- mortgage interest payments (Bank of England 1996: 5).
scientific work. Following this line of thought, the annual percentage change in CPIs ought to be based on a service-flow concept of durable goods, and not on ‘a current acquisition price index that treats durable [good] purchases as instantaneous consumption’ (p. 563).

Yet, it is important to underline at this juncture (as later chapters will show) that, despite appearances to the contrary, any purchase of durable goods is a consumption when considered from an economic perspective. As Verdon has it,

[when an individual buys durable goods, he or she may go on using them for his or her own satisfaction without implying that he or she goes on ‘consuming’. If we want to endow consumption with any meaning at all in an experimental model of complete monetarization we must restrict it to the situation of exchange (Verdon 1996: 162).

Now, as we shall see in Part II, every consumption is an instantaneous action, since it takes an instant (that is, a zero duration in time) to record an exchange, or a payment, in its essential bookkeeping form. And it is necessary to proceed just one analytical step further in order to reach an important finding, whose theoretical consequences might be far-reaching. On reflection, it should indeed be clear that, from an economic standpoint, consumption is nothing more than an expenditure whose object is a pre-existent good. The physical properties of the good consumed need not concern us here, because we are confronted with the cancellation of a commodity from the economic field, not from the material world. This is not to say, of course, that a newly produced Boeing 777 ceases to exist at the very instant when the bank account of (say) British Airways is debited for its purchase price. What we claim is that this payment literally consumes the exchange value of the Boeing 777, whose existence

---

34 One of the many recommendations of the Advisory Commission to Study the Consumer Price Index (1996) focuses on this methodological issue: 'The price of durables, such as cars, should be converted to a price of annual services, along the same lines as the current treatment of the price of owner-occupied housing' (p. 82).
will then be that of a physical value-in-use until it can serve its specific function properly\textsuperscript{35}.

Overall, the distinction between durable and non-durable goods appears to be ill-founded in economics, because it pertains to the physical domain. It is indeed an arbitrary, or at least a subjective, distinction\textsuperscript{36}, which gives a misleading view of the economic nature of output. The whole traditional analysis of inflation may consequently be weakened by it. Is a given price change in a good with longer physical durability any more inflationary than in a good with less? The question is ill-conceived (some would say it is nonsensical).

**From consumer price indices to cost-of-living indices**

The methodological obstacles facing the calculation of the CPI as an accurate measure for the movement in current aggregate prices have led a number of prominent economists as well as various government agencies and private institutions (e.g. trade unions and consumer associations) to advocate the establishment of a Cost-Of-Living Index (COLI) as the objective in measuring consumer prices. Perhaps the crucial recommendation the Advisory Commission headed by Michael J. Boskin makes to the American official statistical agency concerns this specific point (Advisory Commission to Study the Consumer Price Index 1996: iii).

It will help to start with two rigorous definitions.

\textsuperscript{35} Notice the contrast between the economic and physical nature of the object consumed. For expositional convenience we abstract here from the case where the aircraft is resold in a second-hand market. Second-hand transactions will be dealt with later on.

\textsuperscript{36} In effect, to mention a rather extreme example, from a logical point of view a television set is not more ‘durable’ than an apple pie, which can be physically consumed over a week (provided that it is stored in the appropriate way) by eating a slice of it per day. Gordon seems to be aware that the distinction between ‘durable’ and ‘non-durable’ goods is ill-founded, for he claims that ‘[i]t is a sign of the times that many goods like sheets and draperies are officially classified as “nondurable” yet actually last longer than many “durable” goods’ (Gordon 1981: 131).
• 'By the expression “cost of living” we mean the monetary value of those consumers' goods which are in fact consumed in the course of a certain period of time by an average family belonging to a given stratum of a population' (Koniüs 1939: 10).

• 'A cost of living index is a comparison of the minimum expenditure required to achieve the same level of well-being (also known as welfare, utility, standard-of-living) across two different sets of prices' (Advisory Commission to Study the Consumer Price Index 1996: 20).

The COLI is the ratio of the minimum costs a representative household must support, in order to be on the same indifference curve under two price situations. This approach escapes the restrictive, and unrealistic, hypothesis of a fixed consumer basket, for any standard of living may formally be reached by various combinations of goods and services. The underlying concept is thus based on equivalent baskets, rather than on identical baskets as for the CPI (Triplett 1975: 21). So far so good.

Yet, there is no need to go any further into the study of the COLI approach to see that, in principle, a separate cost-of-living index could be developed for each and every household based on their actual consumption basket and prices paid (Advisory Commission to Study the Consumer Price Index 1996: 2, n. 2; Pollak 1998: 69). It is indeed widely recognised today, even among political economists and mathematical statisticians, that '[t]he whole of the theory of the cost-of-living index relates to an individual consumer' (Triplett 1975: 65). Now, in practice, price index analysis never deals with the case of a single household. It may be symptomatic of the demise of price level measurement that two outstanding scholars at the Massachusetts Institute of Technology and at the National Bureau of Economic Research have recently admitted

---

37 Within this theoretical framework, it should be recalled that the cost of maintaining a given standard of living can go down while the price of a fixed market basket can go up (Gordon 1992: 7). Moreover, as Fisher and Griliches cogently note, ‘there are [in reality] many true cost-of-living indices, each corresponding to a particular indifference curve’ (Fisher and Griliches 1995: 229, n. 1).
that ‘[w]hen it comes to dealing with groups of consumers, it is less clear what should be done, even in principle’ (Fisher and Griliches 1995: 230).

Neither the CPI nor the COLI is equipped to account for the specific expenditure patterns of different groups of consumers (Advisory Commission to Study the Consumer Price Index 1996: 30). In both approaches, the compilation of aggregate price indices rests on expenditure data reflecting the preferences of some average, or representative, households rather than those of a heterogeneous population (Quah and Vahey 1995: 1132; Advisory Commission to Study the Consumer Price Index 1996: 71; Pollak 1998: 75; Blow and Crawford 1999b: 5-6). Since consumption bundles vary across households, a long-standing controversial issue has been to depict, by a thorough empirical investigation, which types of households experience the largest price increases over a chosen period. As neatly summarised by Robert T. Michael in a famous paper on "Variation across Households in the Rate of Inflation",

[t]he question of whether certain types of households experience systematically or persistently larger or smaller changes in the price of their market basket gets to the heart of the recently intensified social concern about the distributional impact of inflation on various groups in the economy -- the elderly versus the young, the poor versus the wealthy, and so forth (Michael 1979: 33).

It should be clear that a single, and somewhat synthetic, figure (be it obtained by adopting either a CPI or a COLI formula) cannot take into account, even approximately, the variability in households' consumption patterns inherent in modern dynamic economies. Hence the puzzling question (see e.g. Prais (1959: 126), Pollak (1989: 119; 1998: 70-73), and Deaton (1998: 44)): whose cost of living should a price index represent?

Loosely speaking, it is assumed as a general rule that some average concerning a broadly defined group of households can be used as a roughly good estimator of 'the' national inflation rate. But this point is not corroborated by several empirical studies. On purely factual grounds, concern has indeed been expressed over the inability of a
single aggregate price measure to be representative of different forms of consumers' behaviour (see e.g. the report of the Sandilands Committee (1975: 13), and Moore (1990: 275)). It has been alleged, for example, that 'the substantial differences in expenditure shares suggest that in a period of differences in rates of inflation among different commodities, the inflation rate for different age groups may differ' (Boskin and Hurd 1985: 441). Accordingly, the escalation of a number of formal and informal contracts in proportion to the reference price index may be problematic. For instance, in the United States, 'even though retirees are not members of the CPI-W population, the official CPI-W is the index which is used for determining changes in social security payments' (Hagemann 1982: 495). Yet, using a specific subgroup price index, like CPI-W, for escalating (say) retirement pensions might have an automatic, and unwanted, distributional impact across the whole population. The establishment of a separate price index for the elderly has in fact been a matter of concern, at least in the United States, since the early 1980s (see Boskin and Hurd (1985), Bureau of Labor Statistics (1988), and the literature cited therein).

Now, it is very important to be clear on what any aggregate price index whatsoever is supposed to measure, and the purpose(s) for which it is put to practical use, especially considering the dispute on the choice of the commodities' weighting vector to compute a 'social' COLI. When aggregating over a number of representative households, there are indeed two formal methods to calculate the weight of a particular commodity entering the index, as put forward in the classic (1959) paper by

---

38 In the American context, Hagemann introduced the acronym CPI-R for the price index of retirees (Hagemann 1982: 503). More recently, the BLS has created an experimental price index for elderly consumers (CPI-E) -- defined as aged 62 or older --, which is based upon different expenditure weights for the elderly than those for the rest of the population (Bureau of Labor Statistics 1988: 2-10; Berndt et al. 1997: 13). As a matter of fact, 'seniors get special discounts, for example, and their geographic distribution, and other factors might cause the prices they pay to differ slightly from those recorded in the CPI' (Advisory Commission to Study the Consumer Price Index 1996: 71, n. 70).

39 On the concept of social COLIs and related technical issues, see Pollak (1989: 128-152). Since the definition of a social COLI depends on the definition of a social welfare function, it seems unlikely that government statistical agencies will ever compute such an index (Pollak 1998: 70).
Prais. The conventional method consists in constructing a so-called ‘plutocratic aggregate price index’ (Prais 1959: 127; Diewert 1987: 774), which gives each household’s consumption pattern ‘an implicit weight proportional to its total expenditures’ (Nicholson 1975: 540). The plutocratic weights can be obtained by adopting equation (1.6):

\[
\omega_i = \frac{\sum_{h=1}^{m} p_i q_{h,i}}{\sum_{h=1}^{m} E_h},
\]

where \( \omega_i \) is the (plutocratic) weight of commodity \( i \) (for \( i = 1, \ldots, n \))

\( q_{h,i} \) is the quantity of commodity \( i \) purchased by household \( h \) (for \( h = 1, \ldots, m \))

\( E_h = \sum_{i=1}^{n} p_i q_{h,i} \) is total expenditure of household \( h \).

As the functional form of its weights shows, the plutocratic aggregate price index ‘gives a relatively greater weight to the richer households’ (Prais 1959: 127), that is, to households with relatively large actual expenditures\(^{41}\). Conversely, if one aims to give all households equal importance, regardless of their expenditure, then, according to the terminology introduced by Prais (1959), a ‘democratic aggregate price index’ ought to be established, whose weights would simply be obtained from equation (1.7):

\[
\delta_i = \frac{1}{m} \sum_{h=1}^{m} p_i q_{h,i},
\]

where \( \delta_i \) is the (democratic) weight of commodity \( i \) (for \( i = 1, \ldots, n \)).

\(^{40}\) Since it makes little difference in the present discussion, we ignore the superscripts concerning the two price situations (or the two points in time) entering the comparison, because it will ease the exposition to do so.

\(^{41}\) It may be of some help for interested readers to recall that cost-of-living indices are obtained by the same mathematical technique as that used to compute consumer price indices. So, if a Laspeyres price index is the functional form selected, a plutocratic COLI would be arrived at by substituting the plutocratic weights (the \( \omega_i \)'s) for the standard (base-period) quantity vector (the \( q_i \)'s) in equation (1.2).
Thus, in short, in the democratic price index each household counts equally, while in the plutocratic price index every pound of expenditure counts equally (Pollak 1989: 123; Deaton 1998: 42).

Yet, a question ultimately arises which goes right to the heart of the spurious ‘social’ cost-of-living indices: how to choose in such an approach the ‘true’ functional form for measuring inflation accurately over the economy as a whole, that is, across all sectors and all geographic areas; in short, across all economic agents? Once again, as noted above, there is no scientific way to deal with this problem, essentially rooted in a microeconomic research strategy. As far as democratic and plutocratic price indices are concerned, ‘[a]ny choice among them will depend on assessing the balance of political considerations’ (Prais 1959: 131). Perhaps, to conclude, the main point to be underlined here is the distributional issue involved in selecting one of the many aggregate price measures put forward in the vast economic literature. Indeed, households whose actual expenditure rises less than the percentage change estimated by the targeted price index gain from both the indexed programmes of the general government sector and the escalated receipts they might benefit. The economic agents in the opposite situation lose, instead, part of their purchasing power in what may be termed a zero-sum game over the entire economy.

All in all, the main conclusion to be drawn from this methodological framework of inquiry is that estimating the movement in aggregate prices by means of a COLI cannot provide a more accurate measure of inflation than a CPI can do. Useful though they are, both approaches concentrate on the behavioural influence (mainly in redistributive terms) the process of ongoing rise in prices may have upon several subgroups of economic agents. The fundamental idea behind such a methodology rests indeed on an aggregative, microeconomic assumption, which admits that different sets of domestic income holders may experience different rates of inflation over the same period. Yet, as some macroeconomists like Quah and Vahey (1995: 1136) have put it, the undeniable evidence that in each and every national economy a common monetary
base exists across the entire set of economic units may suggest that there is a unique inflationary process (i.e. a single inflation rate) in any given period (see also Wynne (1999: 1-3)). The importance of the preceding discussion of price index analysis for measuring inflation over a national economy needs therefore to be developed at a more fundamental level. This is the central theme of next chapter, to which we now turn in an attempt to provide further analytical evidence in support of the line of argument we have been developing so far.
CHAPTER 2
FROM TECHNICAL BIASES TO ANALYTICAL ISSUES

In general, literature on price level measurement admits that CPI computation might be subject to a number of biases as a measure of the 'true' movement in aggregate prices. The analysis of biases in official inflation statistics is usually dealt with by referring to the objective of estimating the cost of maintaining a given standard of living1, which is assumed to be a better approach to the inflation problem than calculating the overall CPI. However, as we noted in the previous chapter, the COLI notion does not seem to perform better than actual inflation measurement, because it stems from the same conceptual framework and suffers therefore from the same analytical weaknesses. It is the purpose of this chapter to provide a critical assessment of the methodology used in aggregate price measures, moving from an essentially microeconomic insight, in the first section, towards a relatively new macro-theoretical paradigm that we shall introduce in the second section at an intuitive level. Although there exists a variety of aggregate price measures2, here reference will be made to consumer (or retail) price indices only, since these are the main, if not unique, empirical focus of contemporary inflation analysis.

MEASUREMENT PROBLEMS IN AGGREGATE PRICE INDICES

Broadly speaking, critics agree that among the various defects from which the CPI suffers, the most important ones are (a) failure to take full account of substitution between consumption items in households' expenditure patterns, and between shopping outlets, and (b) difficulty in capturing the change of quality of goods and

---


2 For instance: wholesale price indices, producer price indices, and GDP price indices.
services entering the market basket. The latter defect, together with (c) the difficulty in dealing with entirely new products, forms the so-called ‘compositional’ bias — according to the terminology propounded by Cunningham (1996: 12). Before addressing these and other related issues, it may nevertheless be useful to recall that the aggregation of price data into an overall index may also give rise to a formula bias in the CPI, since the functional forms used in price index analysis already differ on purely mathematical grounds (Triplett 1975: 26; Griliches 1990: 188; Advisory Commission to Study the Consumer Price Index 1996: 4; Greenlees 1997: 175). As we have already noted, the same set of disaggregated data may indeed produce several different estimates of aggregate price movements, according to which index-number formula is chosen for that purpose. Now, let alone the choice of the formula\(^3\), a series of imperfections still affect the CPI as an accurate measure of inflation, in spite of the undeniable evidence that ‘[t]he consumer price index is one of the most carefully researched and best executed statistical programs [in modern market economies]’ (Shapiro and Wilcox 1996: 2).

These imperfections in the measurement of aggregate prices, and hence in that of the rate of variation of the latter, may be classified in two categories, according to whether their effect on the measured price level is a priori indeterminate or if they systematically overstate (understate) price increases. Generally speaking, errors resulting from consumers' substitution of goods and outlets over time have an indeterminate effect on the price level (and might thus to some extent offset each other at the aggregate level), because these substitutions may result from different forms of behaviour other than those elicited by the observed price differentials, and at the statistical level they may not be correctly captured by sampling techniques. On the other hand, failure to take full account of new goods entering the market as well as to

\(^3\) In what follows, a Laspeyres price index is the functional form adopted, since this is in fact the case in most OECD countries.
measure accurately quality changes in existing goods, may impart a systematic, upward bias on measured price variations, because of the 'price cycle' followed by any new good and also because quality is generally increasing over time. Let us address these issues in turn.

**The substitution bias**

Historically, the substitution bias has been the first identified problem affecting the measurement of fixed-weight price indices. According to a fair number of recent studies, it can be separated into two distinct effects, namely item substitution and outlet substitution (see e.g. Manser and McDonald (1988), Gordon (1992), Advisory Commission to Study the Consumer Price Index (1996), Cunningham (1996), Shapiro and Wilcox (1996), Wynne and Sigalla (1996), and Diewert (1997)).

**Item substitution**

It is nowadays a well-established fact that

[s]ince consumers will substitute those goods whose prices rise less or fall more for those whose prices rise more or fall less -- and within limits they can do this without reducing their levels of real consumption -- the fixed-weight base CPI overestimates rises in the cost of equivalent market baskets (National Bureau of Economic Research 1961: 52).

This statement is supported by some algebraic manipulations of a Laspeyres price index, as demonstrated for instance by Pollak (1989: 11-13) and, at a less sophisticated level, by Cunningham (1996: 20-21). In plain language, this is tantamount to saying that, by holding the base-period representative basket constant, the fixed-weight overall CPI puts too much weight on items that have become relatively more expensive and too little weight on items that have become relatively less expensive. The commodity substitution bias thus arises from the consumers' behavioural tendency to shift purchases away from goods and services whose prices
increase relatively faster and towards products whose prices increase more slowly or fall. "Maintaining the basket constant for a number of years under these conditions implicitly gives too much weight to the first type of goods and too little to the second type, and therefore imparts an upward bias in the CPI inflation rate" (Fortin 1990: 128).

The empirical magnitude of the item substitution bias depends ultimately on both the extent to which consumers buy a bundle of commodities different from the representative market basket and the extent to which relative prices move over the period in which weights are fixed. There have been several careful estimates of item substitution bias in aggregate price indices, most of them relating to the American overall CPI. Some early empirical works 'suggest a discrepancy of less than one tenth of an index point per year' (Triplett 1975: 26), which tended to increase in proportion to the increase in prices of individual goods (Braithwait 1980: 73). More recently, using 1959-1985 consumption data from the National Income and Product Accounts, Manser and McDonald's *Econometrica* (1988) paper gives an estimate of about 0.18 percent per year for the item substitution bias affecting 101 commodity groups of the American CPI (Manser and McDonald 1988: 909-910). Using the same methodology, Shapiro and Wilcox predict an average item substitution bias of 0.2 percentage points per annum over the period 1996-2006, if no major changes in relative prices occur in this decade in the United States (Shapiro and Wilcox 1996: 17-18). This also seems to represent a plausible average figure for the Advisory Commission headed by Michael J. Boskin, whose estimated range lies between 0.15 and 0.25 percentage points per annum, even for the years following the benchmark revision of the national CPI in January 1998 (Advisory Commission to Study the Consumer Price Index 1996: 64-)

---

4 See the final report of the Advisory Commission to Study the Consumer Price Index (1996: 21-24) for a hypothetical numerical example of commodity substitution bias.

5 It is pretty obvious 'that lengthening the period of time that the Laspeyres weights are held constant considerably increases the size of the bias' (Braithwait 1980: 73).

6 For an overview of the 1998 revision of the American CPI, see Greenlees and Mason (1996: 6-9).
70). Work by Cunningham on the British RPI suggests, instead, an upper bound of around 0.1 percentage points per year, from which he advocates the sensitivity of the item substitution bias to the frequency of revisions. 'The more frequent the weighting revisions, the less the scope for the index to miss shifts in consumption patterns, and thus the less the scope for systematic [item] substitution bias' (Cunningham 1996: 41).

**Outlet substitution**

It has been alleged by many applied economists that '[a] major trend in the twentieth-century marketplace has been the replacement of small independent "mom-and-pop"-style retailers with large retail establishments owned by chains' (Reinsdorf 1993: 227). Since prices at the large self-service stores often appear to be lower than prices at the small points of purchase, consumers may shift their purchases from high-cost to low-cost outlets over time (Shapiro and Wilcox 1996: 33-34; Diewert 1997: 429). Yet, the procedure for incorporating new stores into CPI outlet samples might cause a substantial long-term bias in the aggregate price index. In fact, because of the method used to rotate outlet samples, the periodic survey of the retail stores where agents purchase their goods and services cannot keep track of actual outlet substitution, 'a process of gradual but steady replacement of higher-priced retail establishments by lower-priced entrants' (Reinsdorf 1993: 235). Furthermore, '[s]ince price data are collected within outlets, the shift of consumers to purchases from discounters does not show up as a price decline even though consumers reveal by their purchases that the price decline more than compensates for the potential loss of personal services' (Boskin et al. 1998: 9). Yet, in contrast to the item substitution bias, 'which can

---

7 In the United States, this is the task of the Point-of-Purchase Survey. 'Outlets are chosen and rotated every five years from a Point-of-Purchase Survey, asking consumers where they purchase goods and services, with probabilities of outlet selection proportional to expenditures' (Boskin et al. 1998: 6). See Armknecht et al. (1997: 379-380) for more details.
continue forever as long as there is dispersion in relative prices, outlet substitution bias must end when low-cost retailers capture the entire market’ (Diewert 1998: 51).\(^8\)

The empirical estimates of outlet substitution bias have been quite rare up to now, mainly because of the monumental data requirements to assess it. Building on the original, and highly detailed, work by Reinsdorf (1993) -- who estimated outlet substitution bias to be 0.25 percentage points per annum for certain food and fuel items sold in the United States during the 1980s --, some scholars have developed their own estimates of an average figure of this bias, be it for the American CPI (Advisory Commission to Study the Consumer Price Index 1996: 67-68; Shapiro and Wilcox 1996: 36), the British RPI (Cunningham 1996: 48-53) or other national aggregate price indices (see e.g. the work by Crawford on the Canadian CPI, reviewed by Cunningham (1996: 46-47)). On the assumption that some 40 percent of the overall CPI market basket would be affected by outlet substitution bias\(^9\), Reinsdorf’s estimated magnitude has led the authors of the 1996 ‘Boskin Report’ to figure out a new-outlets effect of \((0.25 \times 0.4 =) 0.1\) percentage points per year for the official American price statistics (Advisory Commission to Study the Consumer Price Index 1996: 67-68). A similar empirical magnitude has been arrived at for the Canadian CPI\(^10\), whereas the results of the scenarios modelled by Cunningham (1996: 50-53) range ultimately from 0.08 to 0.25 percent per annum for the British RPI.

The new-goods bias

‘The new-goods bias results from the inability of bilateral price indexes to take into account the fact that the number of commodities from which consumers can choose is

---

8 Diewert notes however that, in the future, ‘new outlet competition will come from discount selling of goods over the Internet’ (Diewert 1998: 51).

9 The item groups supposed to be open to bias include: food and beverages, housing maintenance, household fuels, housefurnishings, apparel commodities, motor fuel, medical care commodities, entertainment commodities, tobacco, personal care (Cunningham 1996: 46, n. 12).

10 See the work surveyed by Cunningham (1996: 46-47).
growing rapidly over time' (Diewert 1997: 430). In dynamic market economies where a large number of goods and services are introduced each year, failure to include, in a timely fashion, new items in the CPI basket may lead to a systematic overestimation of aggregate price increases. As argued by Boskin et al. (1998: 8) and Deaton (1998: 38-39), the new-goods bias is in fact quantitatively important and very difficult to deal with, because there is no way to compute it on a routine basis. As a matter of fact, any newly introduced commodity usually undergoes a product cycle of falling price and rising sales\textsuperscript{11}; later on in its 'life cycle', the good becomes 'mature' and ultimately its price grows more rapidly than the average price of its item category (because there is less opportunity for efficiency gains and the item is supplanted by other (new) products). 'The sequence is easily visualized as a "U"-shaped curve -- the price of any given product relative to the consumer market basket starts high, then goes down, is flat for a while, and then goes back up' (Advisory Commission to Study the Consumer Price Index 1996: 34).

To the extent that the government statistical agency's aggregate price measure delays the incorporation of new items into the index\textsuperscript{12}, the overall CPI will tend to have an upward 'new-introductions bias' -- as it has been termed by Triplett (1993:

\textsuperscript{11} Alfred Marshall was among the first to acknowledge the existence of such a product cycle, as the following passage indicates: 'A new commodity almost always appears at first at something like a scarcity price, and its gradual fall in price can be made to enter year by year into readjustments of the unit of purchasing power' (Marshall 1887: 373). Cf. Hicks (1940: 114), but also Diewert (1998: 53) and Wynne (1999: 11).

\textsuperscript{12} Robert J. Gordon forcefully notes that 'the [American] CPI did not introduce autos until 1940, more than two decades after Ford's "Model T" brought the automobile to the average family. Penicillin entered the CPI in 1951, after it had already experienced a 99 percent decline from its initial price. Air conditioning entered the CPI in 1964, more than a decade after the widespread sale of such products. The pocket calculator entered the CPI in 1978, after it had declined in price about 90 percent from early models introduced in 1970' (Gordon 1992: 9). More recently, video-cassette recorders (VCRs), microwave ovens and personal computers (PCs) were included in the American CPI in 1987, 'a decade or more after they had penetrated the market and their price had fallen 80 percent or more. Cellular telephones won't be included in the U.S. CPI until 1998, despite the fact that there are 47 million U.S. cellular subscribers today' (Boskin \textit{et al.} 1998: 10). In the United Kingdom, VCRs, compact-disc players and microwave ovens were not added to the all-items RPI until 1987 -- 'by which time some 43.5\% of households owned a video-recorder' (Cunningham 1996: 37).
200) --, for it misses the price decline that typically happens in the initial stage of the product cycle. ‘In addition, in a Laspeyres price index with a quantity weight that is fixed over several years, the increasing market significance of successful new goods is not taken fully into account, even if they are included -- in small quantities -- in the price index at a very early stage’ (Deutsche Bundesbank 1998: 56). As pointed out by Shapiro and Wilcox, earlier incorporation of new items into the aggregate price index cannot fix the problem within this framework of inquiry. In their own words,

> [e]arly incorporation of new items into the consumer price index will cause them to be underrepresented in the index because they will not have won a significant share of the market compared with the share that they may attain later in their lifecycle. On the other hand, late incorporation will cause the period of supernormal decline in relative price to be missed entirely. The only way out of this dilemma is to combine explicit modeling of the demand for new items with abandonment of the Laspeyres framework (Shapiro and Wilcox 1996: 27-28).

Given the abundance of new items becoming available each year, empirical assessment of this methodological bias is a very demanding task (primarily because of its voracious data requirements). Put simply, the magnitude of the new-goods bias depends upon the proportion of the households' expenditure on new items and the extent of any price cycle followed by the commodities newly introduced (Cunningham 1996: 32). Trajtenberg (1990) attempts to measure this bias in the case of computed tomography scanners (a highly sophisticated diagnostic technology that produces cross-sectional images of the interior of the body), which have been considered as one of the most remarkable medical innovations of recent times. His findings focus on the initial stages of the life cycle of (these) new items and stress the fact that 'the biases stemming from overlooking the strict new goods case are in all likelihood nil, simply because the quantities of new goods sold at the time of their introduction (during, say, their first year), are usually very small' (Trajtenberg 1990: 24). Using econometric techniques to appraise consumer preferences in the ready-to-eat cereal industry, Hausman concludes, by contrast, that the American price index for cereals 'may be
too high by the order of 25% because it does not account for new cereal brands' (Hausman 1994: 29). He maintains that a bias of this magnitude is worth worrying about, for in the period between 1980 and 1992 approximately 190 new brands of cereals were introduced into a pool of about 160 existing brands (p. 7). In the case of the British RPI, Cunningham assumes instead that a set of new electrical appliances as well as audio-visual products enter the marketplace in January 1987 but are not introduced into the domestic price index until 1991 (in fact, he adds, the date of their introduction can be delayed by several years without affecting the results significantly). By varying assumptions about both the change in the new-goods prices and the share of the item groups taken up, he derives a ‘plausible range’ for new-goods bias running from 0.02 to 0.16 percentage points per annum (Cunningham 1996: 38-39). On the whole, however, it is not yet clear whether these, and other\(^{13}\), specific new-goods bias estimates may be valid at a large aggregate level (such as the all-item CPI), because ‘the scientific basis for making a judgment about the magnitude of the new-items effect is particularly thin’ (Shapiro and Wilcox 1996: 31).

The quality-change bias

In many cases the introduction of a new item into the marketplace is in reality a mere change in (some of) the characteristics and features, that is, quality, of an existing product. The use of a fixed market basket fundamentally implies (the simplifying assumption) that the specification of the representative items does not change over time. Yet, ‘[i]f quality is seen as increasing over time (perhaps due to technological advance), a failure to deal with it will bias the RPI above a notional “true cost of living index”, overstating inflation’ (Cunningham 1996: 15). As with the new-goods

\(^{13}\) See e.g. the work reviewed by Shapiro and Wilcox (1996: 30-32), and Wynne and Sigalla (1996: 67-68). More recently, Blow and Crawford (1999b: 45-56) have carried out research on the UK National Lottery -- launched in November 1994 and not yet included in the RPI --, using revealed preference and non-parametric statistical methods to calculate the lower bound on the reservation price of new goods.

62
bias, this problem is rooted in the very methodological apparatus supporting the research strategy of measuring inflation by means of aggregate price indices.

Quality change poses severe problems for a statistical agency. It is non-mechanical in the sense that there is no way to determine quality change on a routine basis. It is heterogeneous in the sense that each quality change is *sui generis* and, like a child, requires individual attention. It is informationally demanding because it may require vast quantities of data that are expensive to obtain and often do not pass the test of a market transaction. Even though routine procedures are established to handle quality change, in the end quality decisions require the subjective judgment about the extent of quality change, and agencies are reluctant to make subjective judgments (Nordhaus 1998: 61).

The main concern in current research is that overall consumer price indices may underestimate the extent of quality change associated with the introduction and diffusion of both improved production processes and better-quality products, and might thus bias price measurement upward\(^\text{14}\) (see e.g. Armknecht and Weyback (1989), Gordon (1990, 1992), Griliches (1990), Triplett (1990), Berndt and Griliches (1993), Liegy (1993), Shapiro and Wilcox (1996), and Raff and Trajtenberg (1997), among many others). It is no exaggeration to claim, as suggested for instance by the Advisory Commission to Study the Consumer Price Index (1996: 31), that the difficult questions posed by quality change represent the ‘house-to-house combat of price measurement’. There is no unique way to compute an average quality-change bias in

\(^{14}\) ‘Obviously we do not want to count as true price increases those price increases that accompany the introduction of new models of existing products that are superior to the existing goods’ (Wynne and Sigalla 1996: 56). It should also be noted in passing that in a few cases quality seems to be deteriorating over time, ‘such as the disappearance of full service gas stations, and the decline in the quality of in-flight service on some airlines’ (p. 62). ‘Other analysts have pointed to reduced convenience and comfort of air travel, deteriorating quality of higher education, increases in travel time and driver irritation resulting from growing traffic congestion, and widespread declines in the quality of customer service as examples of quality decreases that are not accounted for in the CPI’ (Abraham *et al.* 1998: 33). Nordhaus points out further ‘quality deterioration of existing products, which is the exact counterpart of the appearance of new products. The downside of the automobile and air-travel revolution is the deterioration in rail service; typewriters are perfectly adequate for many tasks but have largely disappeared from stores; as software programs are upgraded, they become so complicated that their use seems to require an advanced degree; many electronic devices have daunting instructions; the self-service revolution has led to the demise of many service-oriented retail businesses; the house calls of the family doctor are a fond memory; and so forth’ (Nordhaus 1998: 65).
aggregate price indices\textsuperscript{15}, or to put it more picturesquely, 'there is no substitute for the equivalent of a ground war: an eclectic case-by-case assessment of individual products' (Shapiro and Wilcox 1996: 40). There is indeed still much disagreement about the proper methods to use, and the proper data to employ, to measure quality change for price index analysis (Greenstein 1997: 331, n. 3). As pointed out by Blow and Crawford, '[n]one of these methods constitutes a rigorous, theoretically consistent approach to how quality affects prices' (Blow and Crawford 1999b: 66).

Most academic research on quality changes can be traced back to Griliches's seminal work first published as a staff paper in the 'Stigler Report' (National Bureau of Economic Research 1961: 173-196). In his econometric analysis of quality change, Griliches applied the hedonic regression technique to provide a hedonic price index for automobiles, that is, a particular price index defined on the characteristics of cars\textsuperscript{16}. Further studies for the automobile industry were then put forward, for instance by Triplett (1969) and Ohta and Griliches (1976), along the same lines. More recent leading research -- such as Gordon's (1990) work on \textit{The Measurement of Durable Goods Prices}, Berndt and Griliches (1993), Nordhaus (1997), and Raff and Trajtenberg (1997) -- continued to investigate hedonic price indices for particular industries, but ultimately found that hedonic techniques are by no means a panacea. For instance, as the case is persuasively made by Alan Greenspan, '[t]he benefits of cellular telephones, and the value they provide in terms of making calls from any

\textsuperscript{15} On the conventional methods used to handle quality changes in existing products, see e.g. Gordon (1992: 27-29), Advisory Commission to Study the Consumer Price Index (1996: 36-37), and Nordhaus (1997: 42).

\textsuperscript{16} The "hedonic", or, using a less value-loaded word, characteristics approach to the construction of price indexes is based on the empirical hypothesis (or research strategy) which asserts that the multitude of models and varieties of a particular commodity can be comprehended in terms of a much smaller number of characteristics or basic attributes of a commodity such as "size", "power", "trim", and "accessories", and that viewing the problem this way will reduce greatly the magnitude of the pure new commodity or "technical change" problem, since most (though not all) new "models" of commodities may be viewed as a new combination of "old" characteristics' (Griliches 1971: 4). On hedonic functions and hedonic indices, see Triplett (1987). On hedonic regressions and quality change, see Gordon (1971: 131-141).
location, cannot be measured from an examination of the attributes of standard telephones' (Greenspan 1997b: 3). Sometimes, there are in fact entirely new goods, with fundamentally different characteristics from their predecessors (albeit they both enter into the same statistical category). It thus appears that

The best procedure for empirical price measurement is to combine the conventional and hedonic methods, taking advantage of the relative strengths of each. To the maximum extent possible, the conventional approach should be used to adjust the prices of different models for discrete options, accessories, and added features. Then the hedonic method should be employed to explain the remaining price difference as a function of basic dimensions or performance characteristics (Gordon 1990: 101).17

Turning to the estimated magnitude of the quality-change effect, Gordon (1990: Table 1.2) claims that the American CPI for the alleged 'durable' goods was flawed by an upward bias of 1.54 percentage points per annum on average over the period 1947-1983. However, since measurement problems appear to have declined over time, for the last decade of his price series (1973-1983) he estimates an average bias of 1.05 percentage points per year. Yet, the methodological approach followed by Gordon is open to criticism. His data block, which he derives from mail order (Sears) catalogues and Consumer Reports (two sources independent of those used to compile official price indices), covers only about half of the weight of the 'durable' goods entering the CPI. He explicitly assumes that the rest of the overall CPI is not affected by quality-change bias. However, 'the evaluation that the rest of the CPI is unbiased represents an extreme one-sided answer to the question as to whether the components of the CPI subject to relatively little research are biased' (Advisory Commission to Study the Consumer Price Index 1996: 32). So, by evaluating the American CPI component by component and extrapolating research on quality-change bias from one item group to another when the groups appear to be related, the authors of the 'Boskin

17 ‘Specifically, hedonic techniques are not able to deal with quality changes that are not easily quantified (such as the handling characteristics of a car, the multitasking ability of a personal computer, or whether an item of clothing is in or out of fashion)' (Wynne and Sigalla 1996: 63).
Report (among whom was Robert J. Gordon himself) estimate an ‘aggregate’ quality-change effect of 0.61 percentage points per year for selected time intervals up to 1996 (Advisory Commission to Study the Consumer Price Index 1996: 31-60 and Table 2). This figure is considerably higher than the mean 0.25 percentage points per annum estimated by Shapiro and Wilcox (1996: 40-41) in an analogous mood. In light of both the different methodological standpoints and the diverging factual results of current academic research, it seems thus reasonable to claim, as emphasised by Cunningham, that ‘any estimates should be treated with considerable caution, given the uncertain theoretical base for the [quality-change] bias’ (Cunningham 1996: 24-25).

All in all, although several recent studies address the issue of measurement biases in actual price indices by performing very careful calculations based on disaggregated price data, product specification and the like, it ought to be recalled that ‘all of these estimates come from studies of product specific microeconomic data, and so lack the generality necessary to help gauge the overall bias in the aggregate index’ (Cecchetti 1996: 2-3). As is the case with the calculation of the general price level itself, the assessment of the biases affecting any price index relies in fact on a series of simplifying assumptions that impinge on the pretended macroeconomic validity of the figure so calculated (Deutsche Bundesbank 1998: 57). Such a conclusion is however disregarded by a number of influential researchers. The authors of the ‘Boskin Report’ are a recent example. They calculate a point estimate for the overall bias in the American CPI, by summing up the individual estimates computed for each of the specific imperfections described above (substitution bias, new-goods bias, and quality-change bias), and arrive at an average, upward total bias of 1.1 percentage points per annum, with a range extending from 0.8 to 1.6 percentage points. ‘While 1.1

---

18 A serious problem that must be addressed when trying to draw firm conclusions about the sign and magnitude of the overall bias concerns possible double-counting of some of the biases. In particular, as Wynne and Sigalla cogently observe, ‘[c]an we simply add together estimates of the quality adjustment bias and the new goods bias, given that the distinction between the two is elusive? Is it...
percentage point may seem to be a small amount in any given year, cumulatively year after year it adds up to a sizable difference, 14% over a dozen years’ (Advisory Commission to Study the Consumer Price Index 1996: 70).

Now, it is certainly possible to pinpoint other methodological defects of actual price-level measurement, such as the ‘aging bias’ that might result from pricing in successive periods housing units that become progressively older (Randolph 1988: 359-362; Fortin 1990: 124), or the use of sellers' price lists which do not reflect present market conditions (e.g. discounting20) (Gordon 1971: 123-131; Triplett 1975: 61-63). Similar problems stem from a microeconomic conceptual framework and are fairly straightforward to identify21. However, without anticipating, the main point to be underlined at this juncture pertains to macroeconomic policy -- and its theoretical foundations -- aiming to achieve economic stability. It is indeed well established that, for reasons that should become apparent later on, zero inflation does not fundamentally correspond to a zero rate of variation in any particular aggregate price

possible that traditional substitution bias and quality adjustment bias are aspects of the same phenomenon? The same question can be raised for outlet-substitution bias [...] how do we disentangle this from other more traditional forms of quality and substitution bias?” (Wynne and Sigalla 1996: 82). As recently stated by Blow and Crawford, ‘the direction of bias in the RPI formula is unknown a priori. [...] There is no theoretical presumption of an upward bias in the RPI formula’ (Blow and Crawford 1999a: 2, 29). Yet, as empirical evidence shows, the overall CPI tends to overstate the increase in prices year by year. This means that over the years price measurement biases are compounded.

19 In a similar vein, Shapiro and Wilcox ‘estimate that there is a 90 percent probability that the total bias in the CPI is greater than 0.6 percentage point per year, and a 90 percent probability that it is less than 1.5 percentage points per year’ (Shapiro and Wilcox 1996: 43). The mean estimate of their empirical analysis occurs at just 1 percentage point per annum, thus roughly endorsing the numerical evaluation made in the ‘Boskin Report’. See also Moulton (1996: Table 1) and Greenspan (1997a: 1) for other recent estimates of the overall bias in the American CPI. Note that simulation and extrapolation from available data for the United Kingdom has led Cunningham to guesstimate a ‘plausible range’ for the overall bias in the actual RPI of 0.35 to 0.8 percentage points per year (Cunningham 1996: 57-58). This gives about half the average figure of the American case and may be ascribed to the more frequent weighting revision in the United Kingdom than in the United States (see Table 1.2).

20 ‘It is well known that very few consumers ever pay the list, or “sticker”, price for a car’ (Wynne and Sigalla 1996: 57).

21 This does not mean that they are easily solved, as a quick survey of the most recent literature clearly shows. In fact, as recently stated by a leading central bank, ‘the methodological difficulties [in the measurement of aggregate prices] cannot be completely overcome’ (Deutsche Bundesbank 1998: 61).
index. For example, as recently stated by Mervyn King in a public lecture given at the
London School of Economics to commemorate the fifth anniversary of the Bank of
England Inflation Report, ‘changes in indirect taxes or commodity prices often affect
the domestic price level but do not in themselves change the underlying rate of
inflation’ (King 1997: 8). As a matter of fact, an increase in indirect taxes is much
likely to lead to an increase in the general level of prices, since the goods and services
subjected to increased taxation become more expensive on the marketplace. This
increase in retail prices and, probably, in the targeted price index has definitely a
redistributive effect across the economy, but it does not affect the purchasing power
existing in the whole economy. In fact, if the purchasing power of a more or less
broad group of consumers who make large use of the taxed goods (e.g. fuel, tobacco
and alcohol) is affected negatively by the decision of the State to rise indirect taxes,
the public sector obtains exactly that part of national income lost by the private sector.
An analogous zero-sum process can also be observed within the private sector itself,
when firms raise their mark-up in order to increase their share of total income (i.e.
profits). Certainly the resulting increase in retail prices enables firms to raise profits,
ceteris paribus, but this does not have an inflationary impact on the purchasing power
existing in the economy as a whole, because firms obtain a fraction of income
previously held by other agents (i.e. households). What is lost by one group of agents
(households) is gained by another group of agents (firms), so that the overall process
is a zero-sum game over the entire economy.

22 The first form of income redistribution is from the private sector to the general government sector, in
the case at hand, but a chain of other distributional effects may of course be observed both between
these two sectors and within the private sector itself. To give only an example, consider the tobacco
tax: if the retail price of a packet of cigarettes increases because of an increased indirect taxation, the
present generation of smokers transfers a greater share of national income to the general government
sector, which can then either redistribute it to particular groups of individuals (pensioners and
disabled people, unemployed, political refugees) or invest it according to the current budgetary
policy needs.
No variation in the targeted price index can therefore be ascribed to inflation without further investigation. Furthermore, there may be inflation even when price indices do not vary between two points in time. As pointed out by Chick (1978) with respect to post-war technical change and inflation data both in the United Kingdom and in the United States, the stability of price indices combined with increasing productivity may hide the fact that inflationary processes are at work. In fact, on the assumption that the mark-up does not vary, technical change leading to a reduction of unitary production costs will elicit a parallel decrease in prices. If prices do not decrease and the targeted price index remains stable, this may be ascribed to inflation, which exerts its depressive effect on money's purchasing power. Thus then, computing the variation in an aggregate price index cannot be deemed sufficient to assess the presence, and the magnitude, of inflation in any national economy. Ontologically speaking, a subtler analysis is required, even in the fictitious case where all measurement biases in aggregate price statistics were assumed away. On reflection it is in fact clear that '[t]here is a subtle distinction between targeting a fixed price level versus targeting a zero inflation rate' (Taylor 1997: 8). However, it is fundamentally important to note that this distinction cannot be attributed only to the existence of measurement biases, as is generally done in current literature. In this respect, the following citations are quite telling of the fact that non-measurement issues have been unnoticed so far by a large number of observers.

23 See also De Vroey (1984: 382), who refers to a 1980 Cahier du CEPREMAP by Lipietz and Hausmann.

24 According to Pasinetti (1993: 79-81), it is necessary to break down the general dynamics of prices into an inflationary component and a structural component, since the average rate of productivity growth of the entire economic system is usually very different from the rate of productivity growth in any specific sector of the economy. 'Both components have important and pervasive implications, but of a very different nature. The structural component is tied up with problems of attainment of efficiency in each single branch of production. The inflationary component concerns the economic system as a whole' (p. 81).

25 Note that inflation may occur even in a situation where the targeted price index indicates a reduction in prices. In this situation, the reduction in prices is smaller than it would have occurred if inflationary processes had not been at work.
• 'The existence of upward bias in the rate of growth of the CPI suggests that true price stability will correspond to positive measured CPI inflation' (Shapiro and Wilcox 1996: 52).

• 'Price stability [...] may be defined as 1 or 2 percent measured inflation' (Taylor 1997: 1).

Should monetary authorities and national policy makers be content with a rate of growth in the targeted aggregate price index of 1 to 2 percent per year, because the CPI measurement biases imply that this rate ultimately corresponds to effective price stability26? In Greenspan's words, 'policymakers must be cognizant of the shortcomings of our published price indexes to avoid actions based on inaccurate premises that will provoke undesired consequences' (Greenspan 1998: 1). Indeed, inflation (mis)measurement also matters for the economy as a whole, because it has very important implications for both macroeconomic analysis and policy design. To take an oft-cited example, 'an overzealous pursuit of zero measured inflation may inflict unnecessary costs on society if the price indexes overstate the true rate of price increase' (Wynne and Sigalla 1996: 56).

For the sake of exposition it is convenient to begin by trying to address the main conceptual weaknesses of price index analysis in so far as the inflation problem is concerned. Part II shall investigate more deeply the fundamental theoretical issues.

---

26 To give an example, 'owing to statistical uncertainties, the Bundesbank considers the objective of price stability to be broadly achieved if the measured inflation rate is between 0% and 2%' (Deutsche Bundesbank 1998: 58). For the same reason, the ECB's inflation target is to maintain HICP below 2% (European Central Bank 1999: 46).
ANALYTICAL CAUSES OF INFLATION MEASUREMENT PROBLEMS

Back in 1930 Keynes was already well aware that aggregate price analysis is an imperfect procedure to assess inflation, and his concern is worth quoting in its entirety27.

Most of us were brought up to employ such index-numbers as Sauerbeck's or the Economist's much too light-heartedly, and without sufficient warning that, while there might be nothing better available, nevertheless the actual divergences between these indexes and the Purchasing Power of Money might prove to be, if we could calculate them, of very great significance both theoretical and practical (Keynes 1930: 67).

It would be naive to deny that since the publication of Keynes's Treatise on Money the aggregation of microeconomic price data has been considerably improved by several careful and thoughtful investigations. What we claim is that such methodological corrections cannot be sufficient to understand the problem of inflation. A fundamental analysis of both the nature of modern money and of the working of our monetary production economies is necessary for that purpose. In Keynes's own words, we must face those difficulties of our subject 'which depend rather upon reasoning than upon calculation' (Keynes 1909/1983: 64). To begin with, let us attempt to develop a critical appraisal of price index analysis at an intuitively macroeconomic level. It is the purpose of the next two chapters to explore the theoretical framework in greater detail.

The representative market basket and total current supply

Besides all the methodological biases involved in traditional inflation measurement, there is unanimous agreement among economists that any aggregate price index weighs inadequately or omits altogether some important objects of expenditure -- such

---

27 Joan Robinson (1956: 20-24) pointed out some related conceptual problems in measuring the purchasing power of money, but was silent on the link with inflation analysis.
as several public goods\footnote{It can indeed be very difficult to measure prices for the general government supply. 'For example, how does one price the protective services of the armed forces?' (Steindel 1997: 3). See also Reserve Bank of New Zealand (1997: 14-17).} and personal services (see e.g. Keynes (1930: 57-58), but also Desai (1981: 27-28), Fortin (1990: 110, n. 3), and Bryan and Cecchetti (1994: 195-196)). Yet, the problem of the representativeness of selected items has been treated as a pure sampling problem in most academic research on aggregate price indices. In other words, because of the prevailing empirical approach to inflation measurement, economists have focused on microeconomic data requirements for (only) a certain number of varieties of particular products, instead of extending their analysis to the national economy as a whole. In particular, much discussion about aggregate price analysis has relied on the (often hidden) hypothesis that those varieties not selected for price index construction move together with those varieties which are included in the representative basket. Thus, the position of most scholars, best expressed by Triplett, is to ‘assume that price movements for Mercedes and Jaguar, or lobster and steaming clams (none of which is currently priced) can be approximated by price changes of Chevrolet and Volkswagen, or canned tuna and frozen haddock fillets (which are examples of product varieties priced for the CPI “automobile” and “fish” components)’ (Triplett 1975: 64).

Indeed, if one examines the index number problem in a heuristic manner, one infers that the validity of price index analysis is unavoidably restricted (to say the least) to a specific sample of goods and services, and that the selection of consumption items is purposively directed at not jeopardising the concept of aggregate price indices. Recently, Heymann and Leijonhufvud have achieved an elegant critique of the conventional analysis of inflation, which brings to the fore the fundamental weakness of both CPI and COLI compilation -- as we have already noted in the previous section. In their own words, ““[t]he” inflation-rate itself is a construct of statisticians that may have little behavioural relevance to individuals all of whom are
buying baskets of goods that differ from the CPI basket’ (Heymann and Leijonhufvud 1995: 41, n. 2). Stated somewhat differently, this means that there is no economic rationale for assuming that changes in prices of non-selected items are parallel to those of ‘representative’ products. Neither is there any conceptual justification for eliminating certain varieties of a particular good from the inflation measurement procedure.

This worrying conclusion may be highlighted by a general consideration stemming from a macroeconomic line of thought. Granting the force of price index analysis as an adequate measure of the movement in aggregate prices for a given bundle of commodities, we may still maintain that this framework of inquiry does not provide a satisfactory measure of the purchasing power of money over domestic goods. As a matter of fact, an important, but too often neglected, point which emerges from a widely cited paper by Laidler and Parkin is that ‘analysis of the inflationary process must involve the study of the whole economic system and not just of one or two markets in isolation’ (Laidler and Parkin 1975: 796). Yet, it seems almost trivial to note that measuring inflation by means of changes in a particular price level fails to consider the markets for produced goods as a whole, for this very methodology cannot account for the sum of aggregate output.

In *A Treatise on Money*, Keynes rightly pointed out (in a rather complex statement) that ‘the Purchasing Power of Money [is] the power of money to buy the goods and services on the purchase of which for purposes of consumption a given community of individuals expend their money income’ (Keynes 1930: 54). Now, as we shall see later on, money income is essentially defined by total domestic output, for it is the

---

29 It may be worth noting in passing that ‘[t]o the question, of what our representative commodity should be representative, there is no one answer. It will depend upon the object which we have in view’ (Keynes 1909/1983: 96).
result of the monetisation of all costs of production. 'In each period, say every month, the sum of all incomes formed by the employed "factors", contains the whole range of the new goods flowing from production' (Schmitt 1996b: 86). When we get right to the heart of the matter, this amounts to saying that, in each period, the object of the global, or macroeconomic, expenditure of 'a given community of individuals' (that is, the whole set of economic units) is the total current supply of goods and services produced in the same economy. In inflation analysis, a main conceptual defect of price index application might thus be inferred from a simple representation in which money income is defined by the set of domestic output (Figure 2.1).

\[
\text{CPI market basket} \cap \text{total output}
\]

Figure 2.1 The macroeconomic inadequacy of the representative basket

Unless the CPI market basket really subsumes all currently produced goods and services -- a task which is unanimously deemed impossible in any contemporary national economy --, its use for price index analysis cannot logically pretend to assess money's purchasing power (as lucidly noted by Keynes in the passage of his *Treatise* quoted at the beginning of this section).

---

30 At this stage, it may be useful to recall Keynes's proposal 'to mean identically the same thing by the three expressions: (1) the community's money-income; (2) the earnings of the factors of production; and (3) the cost of production' (Keynes 1930: 123). More on this later.

31 When considering the microeconomic foundations of inflation measurement, Fortin argues that the best measure of the aggregate price level 'is the average price level of all transactions carried out in the economy in a given period' (Fortin 1990: 109). He immediately recognises, however, that this index 'is so broad and would be so costly to construct that it does not currently exist in any country' (p. 110). What he does not seem to consider is the analytical relevance of this ideal price index for measuring inflation (and not simply for evaluating aggregate price changes), an issue to which we shall return.
Now, Fortin maintains that '[t]here are indeed two ways to measure the flow-value of goods and services: through consumption and through production. They are not the same, for two reasons. First, we do not consume all our production -- in addition to consumer goods, we make investment goods, and goods for exports to foreign countries. Second, we do not produce everything we consume -- we also import part from abroad' (Fortin 1990: 110). Since they are closely intertwined in macroeconomic analysis, let us try to tackle these two issues together, although in the scope of this chapter we cannot take the matter at any depth.

From a macroeconomic stance, the measure of the newly produced output must logically be the same when it is taken in the flow of total production as when it is taken in the flow of total consumption\(^\text{32}\). Implicit in this viewpoint is the idea that money income can analytically be divided '(1) into the parts which have been earned by the production of consumption-goods and of investment-goods respectively, and (2) into the parts which are expended on consumption-goods and on savings respectively' (Keynes 1930: 134). Applied to the problem we are investigating, this means that from an economic standpoint it is unrealistic to assume that consumption refers to physical destruction (or transformation) only. In the first chapter we already pointed out that the purchase of a long-lived good (say, a car) cancels it from the economic field, at the very moment the payment is entered into its essential bookkeeping form. If indeed, as some economists seem to suggest (see e.g. Gordon (1990)), investment goods are but another name for durable items, then, conceptually, their purchase falls under the category of consumption\(^\text{33}\). If we define instead as

\(^{32}\) As far as (the depreciation of) the purchasing power of money is concerned, this claim might be supported by the evidence that each pound of expenditure on the factor market is essentially identical to any pound of expenditure on the market for produced goods. If money is affected by a pathology (whose origin will be investigated in the third part of this work) that reduces its purchasing power, it need not be added that each money unit ineluctably loses the same fraction of its original value, be it spent on the factor market or on the product market.

\(^{33}\) A fundamental criticism of the ill-founded idea of durability in economics can be found in Chapter 1. It will be recalled here that, 'even in common parlance, we implicitly term a consumer, not so
investment, or capital, goods those goods bought by firms to contribute to production -- stressing the fact that profits are comprised in the proceeds given in remuneration to the 'productive services' in the sense that they are an income that households transfer to firms on the product market (Cencini 1995: 52-53) --, we understand without pushing the analysis any further that the money income spent in purchasing these investment goods has been saved by households (see Part III for analytical elaboration). In sum, whatever their definition may be, investment goods are part of the newly produced output\(^3\). Accordingly, their exclusion from the CPI market basket adds to the drawbacks of conventional inflation measurement, contrary to the assertion (see e.g. Steindel (1997: 3)) that including the acquisition prices of investment goods into the CPI compromises the price index's ability to track inflation accurately.

As suggested above, a further crucial point concerns international trade, and in particular the alleged phenomenon of imported inflation. What is particularly noteworthy here, is again the problematic performance of the CPI in measuring inflation over the whole national economy. As a matter of fact, by ignoring the changes in the prices of exported goods (since by definition only the domestically purchased items can enter the representative market basket), the CPI misses -- time and again -- an important part of the newly produced output. Further, it wrongly includes the prices of various imported goods (such as crude oil and other raw materials), which have nothing to do with measuring the purchasing power of

much someone who uses up commodities as someone who buys them, as a "big spender"' (Verdon 1996: 107).

\(^3\) It is worth noting that sometimes the same item may be either a consumption or an investment good, depending on the functional classification of the purchaser (Chick 1983: 45). As Verdon observes, "[w]hether or not we choose to call a commodity an "investment good" should have nothing to do with the intrinsic properties of this commodity, but with the uses the individual wishes to put it to" (Verdon 1996: 160). For instance, a car bought for personal (or family) use is a consumption good; used to provide taximeter cab services, it is an investment good.

76
domestic money over domestic output\textsuperscript{35}. In this respect, the GDP deflator appears to be, in principle, a better estimator of the purchasing power of money, because it includes the prices of exported goods and excludes the prices of imported goods. Yet, from a methodological standpoint, the GDP price index is not really better than the CPI, because its compilation is based on the same price data for most of the items surveyed for aggregate price analysis (Moulton 1996: 160). 'Hence, the GDP index shares many of the CPI's technical flaws' (Steindel 1997: 3), and therefore cannot track true inflationary movements\textsuperscript{36}. As we shall explore later on, a more fundamental analysis must be sought, one which takes up Keynes's breakthrough that

\textit{[h]uman effort and human consumption are the ultimate matters from which alone economic transactions are capable of deriving any significance; and all other forms of expenditure only acquire importance from their having some relationship, sooner or later, to the effort of producers or to the expenditure of consumers (Keynes 1930: 134).}

The sample of population and total current demand

It is well known that the bundle of commodities used for assessing aggregate price movements is made up following accurate statistical surveys of the goods and services

\textsuperscript{35} We do not deny that a rise in the prices of imported goods can lead to a generalised reduction in the purchasing power of consumers. To that extent, the CPI might offer an accurate reading of the level of households' satisfaction. But it must always be kept in mind that 'inflation is equivalent to a decline in the internal purchasing power of the pound' (House of Commons 1988: 5). To be more precise, 'inflation is a disequilibrium affecting domestic money in its relationship with domestic output. A change in the price of foreign goods does not alter the purchasing power of domestic money, which can only be exerted over domestic output' (Cencini 1996: 34). This is an important point to which we shall return.

\textsuperscript{36} Considering the contrasting behaviour of alternative price indices, Gordon has suggested that any monetary authority 'needs to decide which inflation index it is trying to stabilize, e.g., the GDP deflator, the deflator for Personal Consumption Expenditures (PCE), or the Consumer Price Index' (Gordon 1996: 6). Indeed, the sometimes widely differing rates of change in these price indices may become a political issue, as testified by the following passage taken from a classic survey of inflation theory: 'The Ikeda Government in Japan [...] relied on wholesale price indexes to show that the rapid economic growth of 1960-62 was not inflationary, while the Opposition relied on consumer price indexes and national income deflators to show the reverse' (Bronfenbrenner and Holzman 1963: 598, n. 8). Note that in the 1970s the wholesale price index was renamed the producer price index (PPI) (Gordon 1990: 9, n. 5).
purchased by a 'reference group' of households in some 'reference' year\textsuperscript{37}. Indeed, it is hardly necessary to stress that 'different people buy different commodities at different times' (Griliches 1997: 171). Since the empirical results of CPI compilation are usually applied to different sets of consuming units (e.g. wage-earners and unemployed; high-income and low-income families; pensioners and invalids), and for different purposes (for instance, indexing government spending or private contracts), a number of authors have been investigating what kind of measurement as well as methodological oddities this practice introduces in price index analysis\textsuperscript{38}. In particular, as noted in the previous chapter, 'some have suggested that different groups in the population are likely to experience faster or slower growth in their cost of living than recorded by changes in the CPI' (Advisory Commission to Study the Consumer Price Index 1996: iii). Whilst this redistributive issue is still an open question\textsuperscript{39}, it is unanimously agreed that any aggregate price index does not account for different expenditure patterns of specific groups of economic units (p. 30).

Now, analysis in this framework should go beyond the distributional impact of inflation on particular subgroups in the economy. In particular, at the conceptual level, one may ask if a representative sample of household budgets can logically be used for assessing the purchasing power of national money over domestic output. Let us try to clarify the basic line of the argument with a representation that now should be familiar to the reader (Figure 2.2).

\textsuperscript{37} Generally speaking, the reference (or survey) year is not the same as the year when the specific price index is conventionally set equal to 100 (Fortin 1990: 116). See Silver and Ioannidis (1994) for further discussion on this point.


\textsuperscript{39} See Berndt \textit{et al.} (1997), as well as other recent empirical studies cited therein.
By analogy with earlier discussion, the essential feature of this approach is that the determining factor for assessing the validity of price index analysis as a proxy for inflation measurement has to rely on a macroeconomic line of thought. Put simply, the essence of the argument is to consider whether the sample of population surveyed can account for the total demand exerted over the whole economy. The answer is negative. Useful though they are, consumer expenditure surveys cannot keep track of total demand, because the method of sampling contrasts with a 'holistic' perception of the entire national economy. Intuitively, even the most accurately-built sample of economic units can reflect only part of current macroeconomic expenditure. By definition, any subgroup of income holders has indeed only a fraction of total money income\(^{(40)}\). Whilst one can certainly examine the effect of price changes across different consuming units by estimating both aggregate demand and income elasticities for certain types of households and for given (categories of) products, it is impossible to say anything about the purchasing power of money (over domestic output) when dealing with the inflation problem at a sublevel of the entire economy. This point is

---

\(^{(40)}\) In this respect, within the set of total demand represented in Figure 2.2, it would have been possible to distinguish the demand of firms from the demand of households. However, since to exert a demand it is necessary to have an income, and since original income holders are households (i.e. wage-earners), the distinction would not affect our analysis at this stage (which, let us recall it, is intended to remain within simple, intuitive boundaries). In other words, since profits are formed on the market for produced goods and services, the demand exerted by firms is ultimately a surrogate of the households' one: it is namely exerted with the income transferred, via the mark-up price mechanism, from the latter to the former. The distinction between firms and households will be considered in Parts II and III, where analysis is refined and developed on macro-theoretical grounds.
fundamental, but it has been unnoticed so far by the economics profession. To investigate the extent to which inflation affects the relationship between money and output, i.e. the relationship that defines the purchasing power of money, attention must be paid to the value of the newly produced output that can be purchased with the sum of money income currently available in the whole national economy. Before addressing this question from a strictly theoretical standpoint, let us bring to the fore a closely related important issue, which may be worth considering within the present context because of the methodology inherent in any aggregate price measure.

The problem of aggregating microeconomic data

Basically, the aetiology of defective analysis of total demand by means of any particular sample of population, as well as of total supply by means of any market basket whatsoever, might lie in the dominant paradigm’s conceptualisation of macroeconomics as being the science of aggregating data over both agents and goods\(^1\). Current macroeconomic models suffer indeed from a serious problem, known as the aggregation problem. This issue is by no means novel (see e.g. Klein (1946)), and much has been written on it\(^2\). In its most basic terms, the aggregation problem has two facets. On the one side, by assuming that it is possible to construct a consistent macroeconomic system by aggregating data over both individuals and goods, the maintained method of economic analysis postulates that macroeconomic relations are a *replica* -- on a larger scale -- of observed microeconomic relations. Thus, statements that are valid for the individual are assumed to hold also at the

\(^1\) As testified long ago by Schumpeter, ‘many economists of our own day [...] divide up economic theory into a theory of the individual firm and a macroeconomic theory that is to take care of the relations between aggregate consumption, investment, employment, and so on’ (Schumpeter 1954/1994: 997). Most contemporary economists believe indeed that ‘the distinguishing feature of macroeconomics is that it collects both individuals and goods together to aggregates and bundles, respectively’ (Felderer and Homburg 1992: 11).

\(^2\) See, for instance, Van Daal and Merkies (1984), Lewbel (1989), Stoker (1993), and Hartley (1997), as well as the abundant literature cited therein.
aggregate level or for the economy as a whole. On the other side, according to this approach it is reasonable to analyse the economic behaviour of a population (that is, a group of individuals) by conceiving and modelling the behaviour of a typical or representative agent, and by transferring the results to the aggregate level. This is in fact one of the key assumptions made in perhaps the most widely used method of studying the macroeconomy, which models the structural relationships in a national economy by working out theories of individual behaviour and applying them to study aggregate behaviour. A recent expression of this approach has been summarised by Hartley in his critical work on *The Representative Agent in Macroeconomics* as follows: 'The first step is to write down the problem faced by the microeconomic agent in terms of fundamental parameters. This agent is assumed to be representative, and the solution to this problem is assumed to hold for the macroeconomy' (Hartley 1997: 26).

So, when one investigates the problem of the measurement (and control) of inflation in the national economy, one might be tempted to derive from the aggregated results of a particular consumer expenditure survey -- carried out with a specific sample of 'representative' agents -- a kind of synthetic, and highly hypothetical, consumption behaviour that should hold for the whole economy under scrutiny. The factual relationship observed between a subset of income holders, on the one hand, and a subset of domestic goods and services, on the other hand, is thus assumed to be valid also for the entire economic system, with no misgivings about the very method of aggregation. Clearly, the national economy is considered, and modelled, 'as if' it were formed by a large number of agents, all of whom have exactly the same economic behaviour as the reference group of individuals surveyed by the nation-wide statistical agencies. The methodological biases occurring in the measurement of aggregate prices, described above, may therefore be ultimately the hallmark of the analytical weaknesses of a research strategy that reduces macroeconomic phenomena
to the aggregation — over both agents and goods — of microeconomic magnitudes. Again, this point has been unperceived in the literature on price level measurement.

As a matter of fact, orthodox macroeconomic models often rely on the Marshallian legacy of codifying a representative agent's behaviour, to derive a corresponding behaviour for a large group of economic units, or, ultimately, for the economy as a whole. As has been emphasised by Kirman, '[t]he motivations for the extensive use of the representative agent are the desires to provide microfoundations for aggregate behavior, and also to provide a framework in which equilibria are unique and stable' (Kirman 1992: 121). Hicks is a famous case in point. When considering the characteristics of the conventional method of economic analysis, he claims indeed that it enables us to pass over, with scarcely any transition, from the little problems involved in detailed study of the behaviour of a single firm, or single individual, to the great issues of the prosperity or adversity, even life or death, of a whole economic system. The transition is made by using the simple principle, already familiar to us in statics, that the behaviour of a group of individuals, or group of firms, obeys the same laws as the behaviour of a single unit (Hicks 1946: 245).

Now, if all 'singletons' were identical, one could quite easily admit that the behaviour of a single economic agent can appropriately represent the behaviour of the sum of economic units of the same type. Yet, in most cases, it is utterly misleading to identify the behaviour of the representative agent with the actual behaviour of the sum of agents acting in any national economy whatsoever. First of all, at the conceptual level, it might be very difficult to give a consistent, and operational, definition of what a representative agent is.

The list of unanswered (and, possibly, unanswerable) questions is lengthy. What exactly is a representative agent? What does it mean to be 'representative'? Does 'representative' simply mean 'average', or does it mean something else? In a group of firms or agents with, say, 100 characteristics, how many of these characteristics must be well-reflected by a representative agent? And so on (Hartley 1997: 16).
To take but an example in our specific field of investigation, it is worth pointing out that a number of empirical studies of aggregate price differentials between various categories of economic agents make, implicitly, the heroic assumption that the population within each subgroup of the economy is homogeneous (see e.g. Bureau of Labor Statistics (1988: 3-10), and Berndt et al. (1997: 34-37)). Yet, as persuasively testified by Michael's (1979: 45-46) and Hagemann's (1982: 506-507) work, the observed price differences of market bundles purchased by different demographic groups in the same country are in fact smaller between these groups than within them. To this analysis one needs only add that, for each subgroup, the representativeness of collected data may also turn out to be highly unstable over time, as we have already noted.

Further, an even more important set of arguments has to be considered here, namely that representative-agent models give rise to a series of fundamental pitfalls when they are used to infer social economic behaviour from (the sum of) individual forms of behaviour (Martel 1996: 128-134; Hartley 1997: 17; Dutt 1998: 313-316). We should indeed not lose sight of the fact that the foundations of representative-agent models are the so-called *homo œconomicus* and the marketplace. Individuals' (rational) behaviour is constitutive of the interactions between economic units, and methodological individualism is assumed to be the ultimate factual underpinning of economic activity. Thus, probabilistic uncertainty and rational expectations play a crucial role in current macroeconomic analysis, for they are fundamental in constructing elaborate stochastic models of the economy. Such complex models might be of some help to grasp the (optimising) behaviour of (representative?) economic

---

43 The concept of methodological individualism (first used by Schumpeter (1954/1994: 889)) portrays the working of the entire economic system by relating ultimately all social phenomena to individuals' behaviour. To quote Boland, 'methodological individualism is the view that allows only individuals to be the decision-makers in any explanation of social phenomena' (Boland 1982: 28). Game theory epitomises perhaps the most clear example of this research strategy, since it deals with the agents' optimising behaviour in a general-equilibrium framework of inquiry. See Walliser and Prou (1988: 109-114) for a fairly non-technical discussion of this topic.
agents -- consumers as well as producers -- in an operational setting. However, granting their usefulness for a number of microeconomic investigations, we may still ask whether these analyses can appropriately deal with, and account for, the working of the national economy as a whole (Verdon (1996: 25-29) is rather critical on this point).

In the traditional method of modelling the economic process by means of a system of simultaneous equations, 'each agent in the economy, whether a consumer or a producer, has at least one equation which describes his or her behaviour. The behaviour of the economy as a whole, at the aggregate, macro-level, is built up from the individual equations at the micro-level' (Ormerod 1994: 78). Thus then, current economic analysis does not admit any solution of continuity between the micro- and the macro-level of the economy. In fact, several recent textbooks have made the (often implicit) assumption that micro- and macroeconomics have the same object of inquiry. So much so that a large number of authors emphasise that the key distinction between micro- and macroeconomics concerns the level of aggregation of economic variables. 'In macroeconomics we are interested in the determinants of broad economic aggregates' (Schlicht 1985: 8). One is thus led to ask when a microeconomic magnitude acquires a macroeconomic dimension. Analyses of phenomena concerning a single economic unit have in general been categorised as microeconomics. Analyses of the whole economic system have usually been attributed to macroeconomics. But how should one classify an investigation about, say, a particular industrial sector or a specific subgroup of income holders of a given country? Which number of economic units is (to be) considered as drawing the

---

44 As Boland would say, 'the basis of macroeconomics is the view that it is possible to keep the aggregated quantities in focus. But most important is the view that all of macroeconomic analysis is methodologically and perfectly analogous to microeconomic analysis. [...] We are saying that if microeconomic theory is true, then the nature of the macroview or the aggregated view of the economy cannot be inconsistent with the microview' (Boland 1982: 84-85).
borderline between a micro- and a macroeconomic framework of inquiry\textsuperscript{45}. A possible answer to these questions can be provided if one considers the conceptual distinction between macro- and microeconomic operations as being analogous to the logical distinction between a set and its elements\textsuperscript{46}. This proposition amounts to saying that a single individual's action may prove to be of a macroeconomic nature, that is, it may concern the set of economic agents, whereas there may exist some situations which do not pertain to macroeconomics although they concern the sum of economic agents. Given the scope of this chapter, we cannot take the time to explore these startling facts thoroughly. However, to get the flavour, let us give two quick stylised examples.

Consider first the case where agent W is employed by firm F during period P to carry out some paid work. (Indeed, this is the hallmark of modern economic activity.) When W receives his or her remuneration for the 'productive services' he or she has provided F, a (new) net value has been formed for the economy as a whole. As a matter of fact, W obtains a claim on a bank deposit (+) corresponding to the debt (−) incurred by F towards the banking system which monetises the transaction; and the goods newly produced (+) are the macroeconomic value \textit{par excellence}. All this will be elaborated later. For the present it might help the reader to sketch a numerical example, and suppose that the value of the newly produced output is measured by 100 units of money. Accordingly, the banking system's balance sheet would match W's deposit (+100) with F's debt (−100), and total current output would be increased (+) by 100. On the whole, (+100 −100 +100 =) 100 units of value would have been

\textsuperscript{45} It need not be emphasised that such an approach does not give entire satisfaction to those who adhere to an organic-cum-holistic view of a monetary production economy. See e.g. Vercelli (1991: 234-237), and Bortis (1997: 365-368).

\textsuperscript{46} Formal logic rejects the definition of a set as the sum of its elements. As a matter of fact, addition (i.e. the operator '+') is not a logical operation on the elements of a set, but it applies to the domain of real numbers (91). This is not to deny, of course, that an addition of the elements of a set is possible when these elements are of a numerical nature. But this operation has nothing to do with the definition of the set.
formed in the operation described above. The productive activity of a single agent (W) gives rise to a net value over the entire set of economic units indeed. Stated a little more fully the above argument amounts to saying that production is a macroeconomic phenomenon, and that its result cannot be thoroughly appraised if one merely considers it as an exchange between workers and firms (an issue we shall explore in Chapters 4 and 5).

By contrast, the sum of all fiscal disbursements paid by the private sector to the general government sector over, say, a fiscal year does not define any macroeconomic magnitude, irrespective of both the size of the given community and the amount of income involved. Fiscal transfers simply redistribute money income in a different way than originally allocated by the remuneration of the factors of production. Even if they involve the sum of economic agents existing in a national economy, these transfers do not have a macroeconomic dimension, because they ultimately depict a zero-sum process: income is neither created nor destroyed by them; only (part of) its 'property rights' are transferred across the economy.

On the whole, contrary to current economic theories, it should therefore become intuitively apparent that the nature of macroeconomic phenomena cannot be appraised through the aggregation of microeconomic magnitudes. Another way of reaching the same conclusion is to note, in Kirman's words, that '[t]here is simply no direct relation between individual and collective behavior’ (Kirman 1992: 118), thus echoing Paul A. Samuelson's argument of the fallacy of composition. Reducing the behaviour of a group of different economic units to the behaviour of a somehow

---

47 As writers like Boland have often observed, '[d]emonstrating the dependence of all macroeconomics on microeconomic principles is essential for the fulfillment of the (methodological) individualist requirements of neoclassical economics' (Boland 1982: 80). Moving away from the microeconomic foundations of macroeconomics then implies -- to use Colander's (1996) metaphor -- climbing down the Walrasian mountain and starting again up another mountain. It is the purpose of Parts II and III to explore a relatively new path.

48 'A fallacy in which what is true of a part is, on that account alone, alleged to be true of the whole' (Samuelson quoted in Hartley 1997: 170).
representative standard utility, or profit, maximiser might lead to conclusions which can be either misleading or wrong (Kirman 1992: 117; Janssen 1993: 84-93; Martel 1996: 140-141). 'It is a fact that the use of a representative consumer assumption in most macro work is an illegitimate method of ignoring valid aggregation concerns' (Lewbel 1989: 631).

Now, there is no real need for us to explore all of these issues, although they ought to be carefully considered by anyone interested in modelling the decision-making process of a single agent and transferring it to the aggregate level. For our purposes, it is sufficient to note that both representative-agent models and microfoundational approaches that aim to bypass the consistency problem in aggregation need the representative agent to portray the set of preferences of every single individual in the national economy. As this proves to be impossible in any economic system where millions of diverse agents have heterogeneous preferences as well as different forms of behaviour, representative-agent models are neither a proper method nor a particularly useful means of studying the economy as a whole (Hartley 1997: 3). When we get right to the heart of the matter, we may thus claim that 'the representative agent methodology described above is a gross fallacy of composition which disqualifies any kind of microfoundation from being a logically consistent and complete foundation for macroeconomics' (Martel 1996: 128).

To sum up, this methodology, widely shared as it is, hardly does justice to the nature of a macroeconomic problem such as the one which is dealt with in the present thesis, namely the alteration of the relationship between money and output. In so far as the analysis of inflation is concerned, doing macroeconomics by aggregation (of microeconomic magnitudes) entails a reductionist vision of the real world, for it fails to consider what Ingham (1996: 509) portrays as the 'social' structural conditions for

---

49 As pointed out by Samuelson referring to the fallacy of composition, '[w]hat is true for each is not necessarily true for all; and conversely, what is true for all may be quite false for each individual' (quoted in Hartley 1997: 174).
the existence of money. To be more precise without anticipating, one has to ask whether the method of calculation, and the representativeness, of any price index can account for the variation in the purchasing power of money over total saleable output. As far as the choice of the functional form of the index number is concerned, we noted that the use of the Laspeyres formula has been strongly criticised by mathematical economists as well as by professional statisticians, without so far putting anything better in its place. Figures 2.1 and 2.2 offer two additional points, which to the best of our knowledge no one has ever raised in the study of inflation by means of price index analysis. To this analysis we then have to add the above considerations stemming from the aggregation problem, which is but another way of looking at the same problems as those put to the fore in the two previous subsections.

Ormerod summarises this set of arguments by stating that the macroeconomic principles of a monetary production economy cannot be deduced from simple extrapolation from the behaviour of its individual components: 'The whole is different from the sum of the parts. There is such a thing as society' (Ormerod 1994: 91). Hence, the need for an organic, fully-fledged macro-analysis arises ultimately from the observation that a fundamentally atomistic paradigm cannot appraise the working of the economic system as a whole50. ‘Microeconomic questions require microeconomic theories although macro considerations might contribute to solve them. On the other hand, macroeconomic questions cannot typically be analyzed fruitfully in microeconomic terms’ (Schlicht 1985: 101). To be sure, there is no such thing as microeconomic foundations of macroeconomics51. The conceptual distinction

50 See Verdon (1996: 13-14). In his rigorous attempt to work out a conceptual synthesis of classical and Keynesian political economy in modern terms, Bortis notes that ‘the question of proportions, i.e. part-whole relations, is fundamental, and this implies that society is primary and is more than the sum of its parts’ (Bortis 1997: 8). In the language of this section we can rephrase this argument by quoting the 1980 Nobel prize winner for economics, who as long ago as 1946 pointed out that ‘[t]here is no reason to assume [...] that there is something sacred about a sum’, when examining the economic significance of aggregates (Klein 1946: 310).

51 Everyone familiar with even rudimentary economics should know that ‘[t]he goal of microfoundations is to explain aggregate relationships in terms of individual behavior’ (Hartley
between micro- and macroeconomics is both analytically clear-cut and operational, and may be supported by the argument according to which analogies between micro- and macroeconomic laws could prove to be seriously misleading, if not entirely wrong (Schlicht 1985: 63-64; Caballero 1992: 1291). To put it yet another way, we might claim that a macroeconomic phenomenon is not the sum of the relevant microeconomic data. So, what is particularly noteworthy here is that no aggregate price measure can ever account for the underlying phenomenon of inflation, since any price index is fundamentally grounded on the (ill-founded) research strategy of doing macroeconomics by aggregation.

The kind of result we are looking for is ultimately similar to the demonstration that, to quote Schumpeter, 'it is possible, as we have put it, to introduce money on the ground floor of general economic analysis without adopting the aggregative view' (Schumpeter 1954/1994: 278). To put it clearly, our aim is to show that a truly macroeconomic analysis of (the variation in) the purchasing power of money has to abandon any aggregative view of the national economy, to embrace a modern theory of money that successfully overcomes the dichotomic representation of the working of our monetary economies of production. Our next challenging task is therefore to attempt to appraise traditional inflation analysis in light of this alternative research programme, and to relay it with a monetary theory of production that conforms to the nature of modern money. It is indeed the purpose of the remainder of this thesis to examine the depth and breadth of the inflation problem from a purely theoretical standpoint.

1997: 132). It is sufficient to note here that this line of thought is deeply rooted in methodological individualism, as suggested previously, and that aggregation of microeconomic data is the hallmark of textbooks' macroeconomic models. As has been clearly perceived by Laidler, 'In the process of acquiring market-theoretic micro-foundations, macroeconomics thus lost its separate identity' (Laidler 1993: 28).
PART II

TOWARDS A MACROECONOMIC ANALYSIS
OF INFLATIONARY DISEQUILIBRIA
CHAPTER 3
THE NEOCLASSICAL ANALYSIS OF INFLATION: A CRITICAL APPRAISAL

As writers like Gale have long and often observed, '[e]conomists' perceptions of inflation rest on measurements of the "general price level" and on rates of change of price indexes' (Gale 1981a: 2). These indices have become an important part of economic investigation, and are in fact considered as the operational counterparts of the (never unambiguously defined) concept of the general price level (Alchian and Klein 1973: 173; Friedman 1987: 5; Greenspan 1998: 3). In the first part of this dissertation we set out to show that price index analysis suffers from important methodological weaknesses (more fundamental than technical defects), which make it theoretically inappropriate for measuring inflation over the entire set of economic units, that is, from a macroeconomic vantage point. The focus on price indices in current inflation analysis might indeed give rise to significant errors in monetary research, in theory as well as in policy making. As we shall argue, adherence to the price level idea implies adhering to the traditional dichotomy between money and output, a theoretical approach that appears to be in contrast with the principles of a monetary economy of production. By the same token, economic policy makers may be misguided in policy design if they rely on price level measurement, because the very measures they focus on in reality may not represent what they assume to be measuring (with a high degree of approximation, given the index number problem that we explored in Part I). We leave it to future research to determine to what extent these errors affect negatively the level of well-being of current and future generations of income holders, both taken as a whole and considered at the 'sublevel' of the entire population of a national economy. A line of argument that might be more relevant to our main theme relies in fact on the grounding of price indices (or aggregate price levels) in economic theory (see e.g. Fisher and Shell (1972: ix), and Triplett (1975:}
19)). It is indeed no exaggeration to maintain, as Schumpeter so convincingly noted in his *History of Economic Analysis*, that

index numbers pertain to the province of the statistical technician and their theory should accordingly be part of the theory of statistics, just as is, for example, the theory of sampling. A great part of the work on index numbers was in fact done by statisticians or by economists who cared little for 'economic theory.' For instance, the formula that of all displayed the most indestructible vitality is due to a man who cannot without qualification be called an economist at all, Laspeyres (Schumpeter 1954/1994: 1092-1093).

Elaborating this point, one might argue that the index number approach to the inflation problem, so deeply rooted in current economic analysis, has failed to bring out a thorough understanding of the underlying causes of the ongoing rise in prices. To put the point sharply, one might claim that in established monetary economics there is a manifest lack of an adequate theory of inflation. As the case is forcefully stated by Hudson, '[t]he consequence of this is that we can offer no cure for inflation with the certainty that it will work' (Hudson 1982: 55). This comment is as relevant today as when it was made nearly twenty years ago, although there has since been extensive writing on the subject (which remains highly controversial in current debate). Woodford (1997: 2) has depicted the current state of inflation theory by observing that the traditional approach to the problem of measuring (the movement in) aggregate prices -- and, so the argument goes, inflation -- takes as its starting point the neoclassical quantity theory of money¹, which is ultimately a theory of the general price level (Laidler 1991: 84).

Yet, before addressing neoclassical monetary economics from a critical point of view, a remark is in order. In the remainder of this thesis we will follow Patinkin's terminology, in so far as the term 'neoclassical' is used here 'as a shorthand

---

¹ Keynes made a similar statement in his essay on "The Method of Index Numbers with Special Reference to the Measurement of General Exchange Value". He claimed in fact that '[t]he origin of the method is to be found in the doctrines of the quantity theory of money' (Keynes 1909/1983: 105).
designation for the once widely-accepted body of thought which organized monetary theory around a transactions or cash-balance type of equation, and which then used these equations to validate the classical quantity theory of money' (Patinkin 1956: 96). Similarly, we will use the term 'monetarism' to characterise the 'contemporary incarnation' (Congdon 1978: 3) of the (still) dominant theory, which considers the purchasing power of money as the reciprocal of the general level of prices (see e.g. Fisher (1911/1931)). While one might quarrel with the existence of a unique, and unambiguous, monetarist doctrine (Mayer et al. 1978: i), it is indeed a widely held view -- also among those economists who define themselves as monetarists (without necessarily endorsing the term itself) -- that the basic idea which lies at the roots of monetarism (as an economic tenet) can be traced back to the quantity theory of money (see e.g. Mayer et al. (1978: 1-5), Desai (1981: 15), Kaldor and Trevithick (1981: 1-2), Laidler (1981: 1-5), and Cagan (1987: 493)). In fact, according to two distinguished adherents to neoclassical monetary economics, '[t]hough it has been under persistent challenge, the quantity theory of money has, in one form or another, dominated the literature on inflation for the greater part of the last three hundred years' (Laidler and Parkin 1975: 744). After an initial exploration and discussion of this theory in the different formulations it has received over the last hundred years, this chapter aims to put forward an internal critique of established monetary analysis, in order to call for an adequate theory of modern money and of the working of our monetary production economies.

---

2 In 1968, Karl Brunner coined and first used the expression 'Monetarist approach' to characterise the body of contemporary monetary theory that emphasises the role of the quantity of money in economic activity (Brunner 1968/1990: 391). In the First Wincott Memorial Lecture, delivered at Senate House, University of London, on 16 September 1970, Friedman expressed some dissatisfaction with the label 'monetarism' that had already been attached to the School of thought to which he made several important contributions (Friedman 1970/1991: 1-2).
THE OPTIMUM QUANTITY OF MONEY AND OTHER ISSUES

The neoclassical (quantity) theory of money receives a variety of treatments in economic analysis, owing to the fact that its advocates sometimes follow different lines of reasoning, although they always put forward the same conclusion as far as the relation between the stock of money (or its purchasing power) and the general level of prices is concerned. 'In its traditional format, the quantity theory of money is a body of doctrine concerned with the relationship between the money supply and the general price level' (Vane and Thompson 1979: 25). Among the various, rather crude, formalisations that one could find in the vast literature, the two most general formulations of the quantity theory in monetary economics are the transactions form and the cash-balance form. Despite their common conceptual ground, these formulations stem from two different methodological lines of approach. It is to them that we now turn, so as to be able to explore their 'contemporary incarnation' critically.

The Fisherine quantity equation

As Hicks recognised in November 1934 before the London Economic Club, the most startling issue in monetary economics (still) lies in the preoccupation of mainstream theorists with the so-called quantity equation. 'This equation crops up again and again, and it has all sorts of ingenious little arithmetical tricks performed on it' (Hicks 1935/1967: 62). In all its versions, though to various extents, the quantity equation rests on a rather simple conceptualisation of the circular flow of income, where the famous Clower's dictum ('money buys goods and goods buy money; but goods do not buy goods') is portrayed as being the distinctive feature of any monetary economy. In

---

3 See Clower (1967: 5). The Clower constraint -- as it has been termed, although its author intends it as an axiom -- serves to make money useful in general-equilibrium models of monetary economies. We shall examine its heuristic implications in due course.
this analytical framework, the goods and services sold are purported to move in the opposite direction to that of the flow of money. 'If the money flow is clockwise, the "real" flow is counter-clockwise' (Dean 1965: xii). Thus, paraphrasing Patinkin's (1956: 1) open statement in *Money, Interest, and Prices*, the natural starting point to study the working of a monetary economy seems to be the commodity market, where (relative) exchanges are supposed to take place between two distinct things or stocks: money and output (Figure 3.1).

\[
\text{moneyagent} \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \ quad
exchange, to convert the right side, $\Sigma pq$, into a form $PT$ where $T$ measures the volume of trade, and $P$ is an "index number" expressing the price level at which this trade is carried on' (Fisher 1911/1931: 195). This is where, and how, the correspondence between the general level of prices and its alleged operational counterpart, the price index, enters established monetary analysis, as noted in Part I. Since the canonical formulation put forward by Fisher, whose mathematical form goes back to the seventeenth century[^4], the fundamental equation of exchange has thus been written as equation (3.2):

$$MV = PT. \quad (3.2)$$

This formalisation[^5] is known as the transactions form of the equation of exchange, because ‘[i]n this version [of the quantity equation] the elementary event is a transaction -- an exchange in which one economic actor transfers goods or services or securities to another actor and receives a transfer of money in return’ (Friedman 1987: 5). As we have previously observed, this statement is the axiomatic definition of a monetary exchange, that is, a transaction where monetary assets exchange for non-money goods (Spindt 1985: 178-180). Indeed, probably the most recurrent, and undisputed, assertion in monetary economics literature is epitomised by the following theorem: ‘By monetary trade I mean trade in which a single good plays a distinctive asymmetric role as one side of virtually all transactions’ (Starr 1980: 263).

Now, quite apart from a careful and thorough investigation into the economic nature of so-called 'monetary goods' (that is, the exact components of $M$, entering


[^5]: For the sake of argument we do not introduce here the longer form of the equation of exchange, which divides the stock of money into fiat money ($M$) and bank deposits ($M'$), and becomes: $MV + M'V' = PT$. See Fisher (1911/1931: 48) for the exact meaning of the newly introduced symbols.
each and every form of the equation of exchange\(^6\), an age-long controversy has arisen on the heuristic character of what has since been known as the Fisherine quantity equation. On the one hand, authors like Friedman have often maintained that 'equations (3.1) and (3.2), like the other quantity equations we shall discuss, are intended to be identities -- a special application of double-entry bookkeeping, with each transaction simultaneously recorded on both sides of the equation' (Friedman 1987: 5). Because one agent’s sale is another agent’s purchase, the sums received for output sales, appearing on the right-hand side of the Fisherine equation, must necessarily equal the sums expended by domestic income holders, which -- owing to the idea of relative exchanges (see Figure 3.1) -- must again equal the existing money stock times the ‘average’ turnover of a unit of money (shown on the left-hand side of the equation) (see e.g. Fisher (1911/1931: 15-17) and Dean (1965: xiv), but also Frisch (1983: 219), Bordo (1987: 175-176) and Duck (1993: 2)). Roy Harrod gives a clear illustration of the necessary truth of the equation of exchange, worth quoting in its entirety: ‘The tautological nature of the equation, if this is the right term to use, springs from the fact that the price of a packet [of cigarettes] is defined as being the amount of money that is given for it. What is handed out in payment for the packet is the same as what is received for the packet. The necessary truth of the equation is simply a generalisation of this for all transactions within given limits of time and space’ (Harrod 1969: 154)\(^7\).

Yet, by contrast, several other writers argue that the Fisherine equation is not an accounting identity but an equilibrium condition. Indeed, to quote Schumpeter, ‘Fisher did not say that $MV$ is the same thing as $PT$ or that $MV$ is equal to $PT$ by definition: given values of $M$, $V$, $T$ tend to bring about a determined value of $P$, but they do not simply spell a certain $P$’ (Schumpeter 1954/1994: 1096). Thus, in

\(^6\) On the problems concerning the definition of $M$, see e.g. Schumpeter (1954/1994: 1097-1098). We shall address this important issue in the next chapter.

\(^7\) See also Schmitt (1959: 923-924).
equations (3.1) and (3.2) one has to substitute the sign of identity (≡) with that of equality (=). As is well known, this argument centres on the probable causal relationships between the four variables characterising the quantity equation. As Patinkin (1956: 1) puts it, the equation of exchange (3.2) can be looked upon as a specific proposition about the causal relationships determining the equilibrium price level. In particular, the general level of prices may be seen as the resultant of forces represented by aggregate demand \((MV)\) and aggregate supply \((T)\) of current output. In Fisher's own words, 'the price level is normally the one absolutely passive element in the equation of exchange. It is controlled solely by the other elements and the causes antecedent to them, but exerts no control over them' (Fisher 1911/1931: 172).

This can be formalised by writing \(P = f(M, V, T)\).

Along these lines, one might then assume, as most advocates of this approach do, that both the transactions volume of goods and the velocity of money are exogenously determined by technologies and institutions (e.g. payment technologies, consumers' tastes, and trade unions), which is tantamount to saying that both \(T\) and \(V\) may be taken as constant in the short run\(^8\): \(T = T̄\) and \(V = V̄\), where a bar over a variable denotes its supposed exogeneity. On these assumptions, the Fisherine quantity equation has in fact been converted into a theory of the determination of the general price level, that is to say, the quantity theory of money or, more properly, the quantity-of-money theory of aggregate prices. Indeed, manipulating Fisher's equation of exchange (3.2) yields the following (equation 3.3):

\[ P = \alpha M, \]  

(3.3)

where \(\alpha\) is a constant since it is the algebraic expression of \(\overline{V} / \overline{T}\).

---

\(^8\) In slightly more sophisticated versions of the quantity equation, the velocity of circulation of the money stock is assumed to be a function of the nominal interest rate. See e.g. Duck (1993: 2).
Equation (3.3) states the proportional relationship between the general level of prices and the quantity of money available in the economy, and represents the most common formulation of the quantity theory of money. 'To use the correct arithmetical term, given the conditions of demand for money, its value *varies inversely* as the quantity available, or in other words the "general level of prices" *varies directly* as the quantity of money available' (Robertson 1922/1937: 32). The main conclusion usually drawn from this analysis has been to derive a quantity equation in rate-of-growth form, that could explain the rate of increase in aggregate prices by the rate of change in the money supply. Indeed, differentiating equation (3.3) with regard to time and dividing both sides by $P (= \alpha M)$ gives the following (equation 3.4):

$$\dot{P} = \dot{M}, \quad (3.4)$$

where a dot over a variable represents a proportionate rate of change.

On the whole, in its strictest (some would say naive) form the quantity theory of money claims that the money supply is the main, if not unique, determinant of the price level, and that the rate of growth of the stock of money determines the rate of inflation -- as measured by the variation in the targeted price index -- from which the given economy suffers over the period under examination. If the assumption of a constant volume of transactions ($T = \bar{T}$) is relaxed, equation (3.4) becomes equation (3.5):

---

9 To be true, later developments of the Fisherine quantity equation questioned the fixity of the coefficient $\alpha$. In particular, in his rather technical "Restatement of the Quantity Theory of Money", Maurice Allais claims that 'the coefficient of proportionality is not constant; its value at each moment depends on the past historical development of total outlay' (Allais 1966: 1153).

10 Albeit considered as a constant in the short run, the velocity of circulation ($V$) of the money stock as well as the volume of output ($T$) may actually vary for several reasons, that are nevertheless portrayed as independent of the variation of the quantity of money available in the economy. Fisher (1911/1931: Chapter 5), for example, affirms repeatedly that the causal relationship between $M$ and $P$ as explained by the quantity theory remains true, whatever happens to the other elements entering the equation of exchange.
\[
\dot{P} = \dot{M} - \dot{T},
\]

(3.5)

thus giving rise to the famous Friedmanite proposition that 'inflation is always and everywhere a monetary phenomenon in the sense that it is and can be produced only by a more rapid increase in the quantity of money than in output' (Friedman 1987: 17).

Yet, the transactions form underlying the quantity theory of money (as well as the theory of inflation, as we are told) has been strongly criticised by its opponents. In particular, Hegeland — echoing Hicks's contention quoted above — asserts again and again that '[t]he only thing the equation of exchange demonstrates is some arithmetic relations between the symbols included. Nothing can be said as to causal relationships between the different elements' (Hegeland 1951: 39). As a matter of fact, 'the quantity theory is no explanation of changes in the value of money' (p. 59), for 'the equation of exchange does not determine the purchasing power of money, i.e., \( P \). As mentioned above, this \( P \) is merely a weighted average of all individual prices' (p. 93). All this will be developed later. At this point, it is sufficient to emphasise the criticism voiced by the 1988 Nobel prize winner for economics, who as long ago as 1966 maintained that

[i]t]he link of proportionality at any given time between the quantity of money and the product of the price level and the level of activity (relation 3.3) is a relationship of interdependence, and from the dynamic standpoint of causality can obviously be interpreted in either of two ways. It may mean that the price level is proportional to the ratio of the quantity of money to the level of activity; or it may mean that the quantity of money is proportional to the product of the price level and the level of economic activity (Allais 1966: 1153).

A similar treatment was also given by Dean (1965: xiv-xv), when he observed incidentally that the value of money's velocity of circulation (\( V \)) has always been empirically defined, and calculated, so as to offset precisely the change in transactions

\[ ^{11} \text{In the symbology of the first chapter we would thus rewrite equation (3.5) as follows: } \pi = \dot{M} - \dot{T}. \]

\[ ^{12} \text{See also Hegeland (1951: 164-166) for a restatement of the same criticisms.} \]
volumes \((T)\)^13. Thus, there is some reason to believe that the alleged causality among the four variables put forward by the Fisherine quantity equation might rely on a *petitio principii*, which, if demonstrated, might already raise some legitimate doubts about the validity of the traditional (neoclassical) analysis of inflation.

**The Cambridge quantity equation**

The Fisherine quantity equation moves from a desire to depict the macroeconomic relationship between the general price level and the money stock using an analysis in terms of circular flows. Yet, another strand of neoclassical (monetary) economists have been seeking to derive the demand for money following a microeconomic approach that considers money merely as a stock. The latter group of quantity theorists focus essentially on portfolio-choice analysis, for their main concern is with the individual’s optimal relationship between his monetary (liquid) assets and his stock of other, non-monetary, goods (Friedman 1969: 73-74; Patinkin 1972: 107-108; Niehans 1978: 4-6; Herman 1984: 583-584; Bordo 1987: 176). All sorts of money-in-the-utility-function models have thus been put forth (see e.g. Sidrauski (1967), Grandmont (1983: 16-32), Poterba and Rotemberg (1987), Woodford (1990: 1073-1084), and the literature cited therein), trying to isolate the determinants of the amount of money a typical or representative agent may wish to hold for transactions purposes.

Ignoring the set of problems associated with the aggregation of microeconomic situations and with representative-agent models, that we already discussed in the previous chapter, the demand function for money \((M^d)\) has usually been formally expressed as the demand to hold liquid assets (the so-called ‘cash balances’) in a certain proportion \((k)\) of the value of total transactions \((T)\) as determined by some index of the price level \((P)\) (see e.g. Patinkin (1956: 97), but also Vane and

---

13 This critique was already made by Joan Robinson (1956: 403-404).
Thompson (1979: 28), Frisch (1983: 220), and Friedman (1987: 6). The optimum quantity of money that should exist in the national economy has then been derived by solving the following simple model, whose equilibrium condition is established in the familiar Marshallian way (money demand equals money supply, $M^d$):

\begin{align*}
M^d &= kPT \\
M^s &= \bar{M} \\
M^d &= M^s.
\end{align*}

Now, whilst equations (3.2) and (3.6) stem from two different methodological approaches, which also make different use of stock-flow analysis in monetary economics, the cash-balance form of the equation of exchange -- attributed primarily to the work of Pigou (1917) and Marshall (1923), and popularised by their Cambridge pupils\textsuperscript{14} -- may be viewed as a mere reformulation of the quantity theory of money derived from the Fisherine equation. The four terms in the Cambridge equation (3.6) have indeed the same economic meaning as those in the Fisherine equation (3.2), provided that $k$ defines the average period of time over which households' cash balances can be used to purchase non-monetary goods and that it is numerically equal to the reciprocal of $V$ (Harrod 1969: 156; Macesich 1983: 26-27; Bordo 1987: 175-176; Friedman 1987: 4-6). In fact, Professor Pigou did not claim any essential difference to exist between his formulation and Fisher's. In his own words,

[although] the machinery that I shall suggest in the following pages is quite different from that elaborated by Professor Irving Fisher in his admirable

\textsuperscript{14} See Eshag (1963: 18-25) on this point. Note that Marshall (1923) never employed any algebraic formula to render in mathematical terms his presentation of the quantity theory of money. In one of his most significant passages he claimed indeed that 'the total value of a country's currency, multiplied into the average number of times of its changing hands for business purposes in a year, is of course equal to the total amount of business transacted in that country by direct payments of currency in that year. But this identical statement does not indicate the causes that govern the rapidity of circulation of currency: to discover them we must look to the amounts of purchasing power which the people of that country elect to keep in the form of currency' (Marshall 1923: 43).
Purchasing Power of Money, and, as I think, more convenient, I am not in any sense an 'opponent' of the 'quantity theory' or a hostile critic of Professor Fisher's lucid analysis. He has painted his picture on one plan, and I paint mine on another. But the pictures that we both paint are of the same thing, and the witness of the two, as to what that thing in essentials is, substantially agrees (Pigou 1917: 39).

Yet, the emphasis by the Cambridge School on the proportion of resources that economic agents choose to keep in monetary form (i.e. liquid assets), instead of focusing on the circular flow of national income as Fisher's equation did, turned the attention of most contemporary monetary theorists to changes in the amount of the individuals' cash holding and their consequences in terms of price level changes (see e.g. Pigou (1917: 54) and Patinkin (1956: 40)). The portfolio-choice approach thus became the standard method for both theoretical and empirical investigation in monetary economics, and the determination of the equilibrium level of aggregate prices has since been viewed as a matter of simple equalisation of money supply and money demand, 'by help of merely arithmetic truisms, which largely reduced the economic significance of the argument' (Hegeland 1951: 97). From then on, neoclassical (general-equilibrium) models of monetary economies experienced an unprecedented mushroom growth, in an attempt to integrate monetary theory and price theory by modelling the interactions of multiple agents and commodities and by assuming money as a device for overcoming 'frictions' (also dubbed 'viscosities') in the commodity market (see, among many others, Hicks (1935/1967: 64-68), Brunner and Meltzer (1971: 800-804), Grandmont and Younès (1972: 355-358),

---

15 As pointed out by the joint winner of the 1974 Nobel prize for economics, '[i]t is a peculiarity of all systematic treatises on orthodox economic theory that there is no inner connexion and integration of monetary theory with the central theory of prices. Usually the monetary theory is only a rather loose appendix to the theory of price formation. The central economic problems -- according to the classical theory, those of production, of barter-exchange and of distribution -- are treated, without exception, as problems of exchange value, or in other words as problems of relative prices. Obviously, by regarding the central economic problems in this way one entirely detaches their fundamental treatment from any monetary considerations' (Myrdal 1939: 10).
Clower (1977: 210-211), Starr (1989: 3-6), as well as a number of seminal contributions included in the latter volume).

The modern quantity equation and the expected inflation rate

The modern version of the quantity theory of money emanates from an attempt to circumvent what his archpriest -- to use Kaldor's (1970: 2) term -- labelled 'the ambiguities of the concepts of “transactions” and the “general price level”', that, as he lucidly observed, 'have never been satisfactorily resolved' in received economic analysis (Friedman 1987: 5). In fact, the large amount of research work carried out by quantity theorists over the last three hundred years has established neither a fully-fledged measure of transactions nor a comprehensive price index related to them (Bordo 1987: 175-176). As a result, a shift has occurred in monetary thought from the transactions version of the quantity theory of money to its income version\(^\text{16}\). Since the 'oral quantity theory tradition' developed at the University of Chicago throughout the 1930s and 1940s\(^\text{17}\), monetary economists have tended to make use of the quantity equation in its modern income form, which metamorphoses equations (3.2) and (3.6) into the following expression (equation 3.9) -- viewed by Friedman (1987: 6) as a halfway house between the Cambridge version and Fisher's:

\[
MV = Y, \quad (3.9)
\]

\(^{16}\) See Howells (1996: 107-111) for an interesting discussion of these two versions in connection also with the endogenous money approach. In Part III we shall address non-income related transactions and their relevance to money emission.

\(^{17}\) Patinkin (1969) questions the existence of a distinctive Chicago tradition, as well as its analytical link to Friedman's monetary economics, which he depicts as an extension of Keynesian liquidity preference theory. With the publication of Patinkin's (1969) article, the notion that the Chicago oral quantity theory tradition was an invention spread rapidly and became widely accepted indeed, as testified by a number of historians of economic thought quoted by Tavlas (1998: 212-213). However, as argued extensively by Tavlas (1997, 1998), there really existed a Chicago oral tradition during the early 1930s, which was not to be found outside of Chicago and which contained important links to Friedman's monetary analysis, culminating in his restatement of the quantity theory. On the relationships between Friedman's thinking and Keynes's, see Dostaler (1998).
where \( Y \) represents current national income (see e.g. Friedman (1959/1969: 139; 1971: 329), Patinkin (1972: 95), and Bordo (1987: 175)). Hence, the necessary and sufficient condition for income growth (as well as for inflation, according to this postulate) has been written as follows (expression 3.10):

\[
\dot{Y} \leftrightarrow \dot{M} \leftrightarrow \dot{P},  
\tag{3.10}
\]

so that the term in the middle is the \textit{sine qua non} condition for the left- and right-hand terms (Stein 1976: 254; Mayer et al. 1978: 5; Laidler 1981: 1).

Now, it has been claimed that '[t]he modern quantity theory, and the monetarist school based on it, makes several fundamental departures from the neo-classical quantity theory' (Johnson in Mayer et al. 1978: 131). In particular, whereas in the 'old' quantity theory the causal link ran from money to prices (see especially equations 3.3, 3.4, and 3.5), in the 'new' quantity theory the direction of causation is from \( M \) to \( Y \) (on the assumption that \( V \) depends on a set of variables none of which is affected by (changes in) the money supply). The essence of monetarism is indeed that changes in the money supply cause proportional changes in national income, after a variable time lag (see e.g. Friedman's (1968) Presidential address delivered at the Eightieth Annual Meeting of the American Economic Association). The relevant passage in Friedman's dictionary article is worth quoting in its entirety: 'Today's income growth is not closely related to today's monetary growth; it depends on what has been happening to money in the past. What happens to money today affects what is going to happen to income in the future' (Friedman 1987: 16)\(^{18}\).

Yet, without going much further for the time being, on closer investigation the alleged explanatory power of the monetarist doctrine seems even more questionable than that of the 'old' quantity theory, in so far as the effect of monetary expansion on aggregate prices is concerned. Paul Davidson and Sidney Weintraub correctly note

\^{18} See also Hume (1752/1955: 37-40).
that, if national income is affected by money supply growth as monetarists claim, 'how much of the money increase runs in the form of price movements, and how much in output and employment, is left obscure' (Davidson and Weintraub 1973: 1120). Indeed, the consolidation of \( P \) and \( T \) (see equation 3.2) into a unique variable \( (Y) \) as in expression (3.9) entails *ipso facto* the annihilation of the distinction between price movements and transactions (that is, output) fluctuations, and the whole quantity-of-money theory of aggregate prices is consequently obscured by this fact.

To be true, Friedman himself explicitly acknowledged this problem of the modern quantity equation, when he observed that the latter 'has nothing to say directly about the division of changes in nominal income between prices and quantity' (Friedman 1971: 337). He maintained nevertheless what might fairly be called the 'Friedman Rule' regarding the optimum quantity of money, which states that the best monetary policy to provide a Pareto-efficient allocation of resources, and thus maximise the level of well-being of (representative) agents, would be to maintain a rate of growth of the money supply so that 'the quantity of money per unit of output can be kept from increasing appreciably' (Friedman 1987: 17). The idea behind such a quantity-theoretic proposition has been popularised by the policy-oriented assertion that the rate of money supply growth -- minus the growth rate of output, as in expression (3.5) -- should be kept constant over the long run, so as to produce a constant (that is, predictable) inflation rate, defined as the rate of growth of the targeted price index.

As Frisch puts it, '[t]he recent reformulation of the quantity theory is the accelerations theorem. It implies that only an acceleration or a deceleration of the rate of money growth produces any real effects, i.e. employment and output effects, while a constant rate of growth of the quantity of money determines the rate of inflation' (Frisch in Mayer *et al.* 1978: 114).19

---

19 See also Brunner and Meltzer (1976: 154-155).
So, granting the force of this conception, the ‘rational-expectations revolution’ of the 1970s shifted the attention from the causal link between money supply growth and price changes to the distinction between expected and unexpected inflation (Sargent 1996: 539). In fact, the unpredictability of inflation, that is, uncertainty, has become one of the centrepieces of contemporary monetary economics, inasmuch as Friedman’s 1976 Nobel lecture highlighted that ‘what matters is not inflation per se but unanticipated inflation’ (Friedman 1977: 458). Indeed, the rational-expectations hypothesis claims that the individuals’ economic rationality enables them to use both currently available information about the variables of concern to them and a correct theory of the interrelationships among those variables to anticipate future variables adopting optimal forecasting methods (Barro 1976: 15-25; Friedman 1987: 14-15). Hence, any anticipated growth of the money supply, like those ensuing from proclaimed governmental policies deemed credible by the private sector, would be immediately incorporated into inflationary expectations -- via the monetarist ‘transmission mechanism’ --, so that the resulting movement in aggregate prices would have ultimately no ‘real’ effect on national income (i.e. \( Y \), the consolidated variable in equation 3.9). To quote Stanley Fischer on this point, ‘any anticipated monetary policy action will not affect output. Rather, such actions are reflected in both the expected and the actual price levels, leading to no effect on output. The result [...] is that monetary policy actions affect output only if they are unanticipated - meaning not reflected in pricing decisions’ (Fischer 1987: 648).

A vast literature on this forward-looking approach to inflation\(^{20}\) and its effects on output has then spawned neoclassical models of money-using economies (see e.g. Lucas (1972: 119-122), Sargent and Wallace (1975: 242, 246-249), Fischer (1980: 211-220), McCallum (1980: 717-718), and Taylor (1985: 392-396, 403)). So much so that, within monetary macroeconomics, Walrasian rational-expectations

\(^{20}\) Contrast the backward-looking approach encompassed by price index analysis, as examined in Part I.
equilibrium models represent, today, the canonical way of dealing with individuals' maximisation problems in a consistently understood environment (Sargent 1996: 545)\textsuperscript{21}. In particular, over the last twenty years or so, neoclassical monetary economists have been spending much time in order to build so-called 'overlapping-generations models', trying to cast new light on the existence of monetary (dis)equilibria in a framework revived by various sorts of game-theoretic approaches in the rules-versus-discretion debate. These models of money-using economies generally assume a two-period lifetime for each and every generation of agents\textsuperscript{22}, in order to gauge inflation forecasts under the rational-expectation hypothesis of finitely-lived consumers. The \textit{locus classicus} of all such analyses is Samuelson's 1958 \textit{Journal of Political Economy} article, which put forth "An Exact Consumption-Loan Model of Interest with or without the Social Contrivance of Money". The approach suggested in this seminal paper stems from the research strategy according to which overlapping generations of consumers represent a sort of trading viscosity, something that inhibits the sacred Walrasian market clearing at each point in time. It has then been ambitiously claimed by one of the most outstanding proponents of this approach that 'models built on this [overlapping-generations] friction can be made to confront virtually every long-standing problem in monetary economics' (Wallace 1980: 51)\textsuperscript{23}.

\textsuperscript{21} This approach is also known as the 'new classical macroeconomics' (see e.g. Fischer (1987)). Yet, one might object that '[t]he assumption of rational expectations which presupposes the correct understanding of the workings of the economy by all economic agents -- the trade unionists, the ordinary employer, or even the ordinary housewife -- to a degree which is beyond the grasp of professional economists is not science, nor even moral philosophy, but at best a branch of metaphysics' (Kaldor and Trevithick 1981: 15). More recently, Laidler forcefully expressed the same concern in his 1988 Presidential address to the Canadian Economics Association, when he stressed that 'it is dangerous uncritically to model agents' behaviour on the assumption that they form their expectations using the same model of their economy as the economist who studies them. [...] [A]ttempts to model agents' learning processes "as if" they were econometricians seeking information about the economy by statistical induction seem to me to be of limited usefulness' (Laidler 1988: 702-703).

\textsuperscript{22} Woodford (1987) set forth an overlapping-generations model where all agents live for three periods.

\textsuperscript{23} It might help the reader to recall that the frictionless (non-monetary) Arrow-Debreu general-equilibrium model denies the existence of money as a means of payment or as a 'temporary abode of purchasing power', to use Friedman's consecrated phrase, 'basically because the Walrasian auctioneer obviates any need for a medium of exchange' (Fuerst 1994: 582, n. 2). To account for the
To be true, advocates of this postulate mainly follow the cash-balance form of the quantity-theoretic analysis of monetary economies, because they consider money as a distinctive, intrinsically useless (if not worthless) asset that is known to be acceptable to all generations, whose utility functions always include its transaction-facilitating services\(^2^4\) (see e.g. Lucas (1972: 104-109; 1980: 136-140), Wallace (1980: 52-60), Balasko and Shell (1981: 114-121), Sargent and Wallace (1982: 1215-1227), McCallum (1987: 328-329), Woodford (1987: 53-63), but also Bernhardt and Engineer (1994: 495-499), Duffy (1994: 543-548), Bertocchi and Wang (1995: 208-210), and Gottardi (1996: 78-80))\(^2^5\). The following passage -- by one of the most lucid critics of this method of monetary investigation -- will serve to illustrate the crucial assumption made in all kinds of overlapping-generations models of monetary production economies: 'Workers can acquire [the monetary asset] by saving when young and then sell it for consumption goods when old. It is tempting to call this asset \textit{money} and to exclaim, Eureka, here is the reason for the existence and value of \textit{money}' (Tobin 1980: 83). Put differently, in such a setting the mythical, and boldly mystified, double coincidence of wants feared by all neoclassical economists is reported to be ingeniously overcome by introducing a durable good called money, and further assuming that -- in each market session -- any purchase of (consumption) goods by the 'old' generation is a demand for labour (i.e. output) to be provided by the 'young' generation.

existence of money in neoclassical analysis one has therefore to assume \textit{nolens volens} some kind of friction that hinders the clearing of all markets instantaneously and that money can help alleviate. See Starr (1989: 3-6) for a concise discussion of the integration of money in general-equilibrium theory, and Geanakoplos (1987: 771-777) for a summary presentation of the overlapping-generations method of monetary investigation.

\(^2^4\) Within modern neoclassical theory money is consequently treated like a perfectly durable good, since its service-flows enter wealth-holders' utility functions in exactly the same way as it is suggested be the case for other categories of durable commodities such as cars and owner-occupied housing. Yet, in the first chapter of this work we already noted that the distinction between durable and non-durable goods is ill-founded in economics. This also raises the question of treating money as a peculiar good, an issue that we shall investigate in the next chapter.

\(^2^5\) Surveys of overlapping-generations monetary models are provided by McCallum (1990: 981-985) and Woodford (1990: 1106-1116). See also Geanakoplos's (1987) dictionary article.
There is no real need for us to explore in detail all of these models, whose common
basic idea relies on the concept of relative exchange schematically represented in
Figure 3.1 (and which will be taken up again in Chapter 4). The existing differences
between rational-expectations, infinite-horizon models and overlapping-generations
models with finitely-lived agents (the two most used contemporary approaches to
study the working of monetary economies) are indeed more a matter of emphasis than
substance. They all share the belief – unanimously held by neoclassical economists –
in a dichotomous world, where money is introduced in an Arrow-Debreu (general-
equilibrium) framework of inquiry to accommodate some portrayed frictions in factor
as well as in goods markets. To this analytical vision we now turn, in an attempt to
mine from inside the fundamental pillar of both the traditional quantity theory of
money and its contemporary incarnation(s).

THE DICHOTOMOUS REPRESENTATION OF MONETARY ECONOMIES

Over the years, as pointed out by Frank H. Hahn (1982: 21), the picture behind the
now so-called 'cash-in-advance constraint' has axiomatically postulated what
neoclassical theorists have been seeking to account for in their research programme,
namely both the existence and the value of money. A large amount of theoretical as
well as empirical research along the lines of the most influential paradigm in

\[\text{Note, however, that overlapping-generations models can be an appropriate research strategy in some areas outside monetary macroeconomics. In particular, as has been suggested by Baranzini (1991: 89-91) and Pasinetti (1993: 109, n. 6), the overlapping-generations analysis might be an appropriate framework for those theoretical investigations concerning individuals' decisions on how to distribute consumption and savings over their lifetime and with respect to inter-generational bequests.}\]

\[\text{See e.g. Lucas and Stokey (1983: 78-82), whose model introduces the distinction between 'credit goods' and 'cash goods', so that the former can be purchased with current money-income but the latter only with (fiat) money previously accumulated. The first model to put forward such a constraint may easily be traced back to Clower's (1967) classic paper, where the author set to express analytically the requirement that money must be held before an exchange is to take place, for -- as Clower's most famous phrase states -- goods do not buy goods. See Fuerst (1994: 582-585) and Huo (1995: 833) for more recent restatements of the cash-in-advance constraint along the lines put forward by Robert E. Lucas and Nancy L. Stokey.}\]
(monetary) macroeconomics has thus developed and refined the dominant view of the workings of contemporary production economies, where the banking nature of money could no longer be denied. Such an approach has long been trying to fit in an Arrow-Debreu fictional economy some kind of commodity- or fiat money, whose purported essential aim is to function as a means of exchange as lubricant as possible (see e.g. Hicks (1966/1967: 2-7) and Hahn (1973: 230-234; 1987: 21-29), but also Rogers (1989: 58-67 for a thorough critical appraisal). Yet, an acute observer of the history of economic thought has recently noted that

the world of GET [standing for general-equilibrium theory] is in fact a dream world, a world which is not totally workable in the context of actual society. The number of actors on the stage in this GET world are far too few. [...] I will grant that such a retrogressive approach is an easy path if one wants to construct a model axiomatically, and it may also be of some use as a temporary means of facilitating the dichotomizing method, whereby entrepreneurs and bankers are put to sleep for a while (Morishima 1992: 198-199).

As a matter of fact, it is a widely-held view, even today, that our economic systems are characterised by two dichotomous spheres: the real sphere, where output and relative prices are determined, and the monetary sphere, whose unique function is to establish -- via the best-fitted quantity equation -- the general level of prices (and hence current national income) by means of the interaction between money supply and money demand. In this analytical framework, Harry G. Johnson maintains for example that

[the classical quantity theory, in its equation of exchange formulation, had the useful and necessary purpose of separating the theory of real equilibrium (relative prices and quantities) as determined by factor quantities, technology, and preferences, from the determination of money wages and prices -- establishing what came to be known as 'the classical dichotomy' or 'the neutrality of money' or 'the homogeneity postulate' (Johnson in Mayer et al. 1978: 127-128).]

Bank money has been analysed by Post-Keynesian monetary economists, as we have shown elsewhere (see Rossi 1998, 1999).
This quotation has the merit of summarising in a single statement the threefold hallmark of received monetary theory, whose principal tenets are the neutrality of money and the homogeneity postulate, which are ultimately but a corollary of the traditional dichotomy referred to above. Though the depth and breadth of this familiar ground have already been well explored, a fresh look at it will serve us as a Trojan Horse to pinpoint some unfamiliar, and probably unpleasant, monetarist (or quantity-theoretic) pitfalls.

**Money and output as two distinct objects**

Let us reconsider the numerical example introduced in the first chapter, where we assumed that in the given national economy a money stock of 2,000 macro-units faces (or, as we are told, exchanges with) 100 macro-units of output, thus giving rise to a general price level of \( P_1 = 20 \). Analysed more closely, the paradigm underlying this theory raises serious doubts. There is in fact a fundamental oddity in the very definition of the general price level, as it so often appears in recent monetary economics literature. This problem arises out of the fact that the stock of goods cannot possibly be measured independently of money. Indeed, by relating the money stock with the stock of output the given quantity of money will buy, one does not explain where the measure of output comes from. In other words, if money and output really were two distinct things, whose asserted confrontation on the market ring would enable one to determine the current level of aggregate prices, how would it (ontologically) be possible to measure output without having recourse to the monetary sphere? How can output be valued -- not to say exist -- as an economic object, if it is abstracted from money (by considering the latter as a stock existing independently of the former, and vice versa)? And again, how can we posit in the above fictitious example that total output is equal to 100, without considering (in a way that shall be further investigated in the following chapter) what neoclassical theorists call the
'dichotomous' monetary sector? At first glance, one might claim that output is measured in mere physical terms, so that in our economy there are, say, 100 tons of beef which encompass total national product of period \((t = 1)\). But it should take only a moment’s reflection to see that this answer hardly does justice to the complexity of sophisticated market economies like those of the real world. It seems in fact almost trivial to note that any advanced economy currently produces multiple goods and services. In economic terms, their measure cannot be established unless they all have been homogenised by their monetary form, whose numerical magnitude measures national output objectively. Is it possible to add up, say, two tons of Scottish smoked salmon and five tons of Irish roast beef without resorting to their monetary measure (expressed by a number of pounds, or perhaps euro, as we shall see later)?

To concentrate attention on the problem we are investigating, it should be stressed that the attempt to explain inflation by observing the variation in the level of prices -- by means of any index-number formula -- openly contradicts logic, as soon as the intimately related requirement of measuring output objectively is brought to the fore. If, as they do, neoclassical economists define the general price level by the relation they virtually establish between the stock of money and the stock of output (and then consider any variation in that level as the reciprocal variation in money's purchasing power), their research programme on inflation is seriously weakened by the axiomatic dichotomy they establish between these two stocks. In fact, it is on this specific, but fundamental, point that mainstream economics goes astray, as has been suggested by one of the most distinguished scholars of neoclassical thought. In his endeavour to provide *A New Formulation of General Equilibrium Theory*, as the subtitle of his 1992 book indicates, Morishima recognises indeed that 'the method of analysis dichotomizing economics into two specialized departments, real and monetary, is harmful and defective. We must deal with the economy as a whole uniting and interlinking the two subsystems' (Morishima 1992: 184).
Now, without anticipating, one may still argue that central to all versions of the neoclassical (quantity) theory of money-using economies 'is a distinction between the nominal quantity of money and the real quantity of money' (Friedman 1987: 3). It is in fact a well-known commonplace of received literature that '[t]he monetary analysis outlined [in it] differentiates explicitly between real magnitudes and nominal magnitudes' (Brunner 1970: 24). The nominal quantity of money is a nominal magnitude in the sense that it is the number of money units which constitute what has been termed the money stock \((M)\), as captured by the widely used statistical definitions adopted by central banks and international organisations around the world \((M_0, M_1, M_2, M_3, M_4\) etc.). The real quantity of money is 'the volume of goods and services the money will purchase' (Friedman 1987: 4); it is, in other words, the real object (or the purchasing power) of nominal money, and in this research strategy it is precisely obtained by dividing \(M\) by the price level in the light of the analysis put forth by the Classics, particularly in Smith's *Wealth of Nations*, this line of reasoning could have been developed so as to firmly establish why its author claimed that 'the wealth or revenue [...] is equal only to one of the two values which are thus intimated somewhat ambiguously by the same word, [...] to the money's worth more properly than to the money' (Smith 1776/1970: 386). In fact, an investigation into the nature of modern money might show that the distinction between money proper and money's worth, that is, the distinction between nominal and real money, might be considered as being analogous to the distinction between a numerical form and its real content (an issue we shall explore later on). However, since in neoclassical analysis the link between the nominal and the real quantity of money is provided by the price level, which is theorised as resulting from the confrontation of the stock of money with the

\[m = M + P,\]

Recalling that \(P = M / \Omega\) (see Chapter 1) and rearranging the terms, this is tantamount to saying that \(m = \Omega\).

---

29 The real quantity of money \((m)\) may thus be written as follows: \(m = M + P\). Recalling that \(P = M / \Omega\) (see Chapter 1) and rearranging the terms, this is tantamount to saying that \(m = \Omega\).
stock of output, the investigation referred to above has simply been eschewed or, to borrow Morishima's phraseology, put to sleep for a (rather long) while.

The vicious circle inherent in received monetary analysis should by now be evident and may ultimately be formalised as follows:

\[
\frac{M}{P} \Rightarrow m
\]

and

\[
\frac{M}{Q} \Rightarrow P,
\]

where the direction of the arrows indicates causality, and \( m = Q \) for the reason stated above. Thus then, the general level of prices is seen as determining and, at the same time, as being determined by the relationship existing between (nominal) money and (real) output. There is then some reason to believe that such crucial issues in monetary economics need re-examining. One might raise indeed two closely related questions, to which we now turn.

Money, a veil?

Another argument often put forward by neoclassical theorists to provide analytical foundations to their dichotomous view of our economic systems refers to a romantic interpretation of the Pigovian idea of the veil of money\(^\text{30}\). For the sake of analysis it is worth quoting its original source at some length on this point. 'Take the real facts and happenings away, and the monetary facts and happenings necessarily vanish with them; but take money away and, whatever else might follow, economic life would not become meaningless: there is nothing absurd about the conception of a self-sufficing family, or village group, without any money at all. In this sense money clearly is a veil. It does not comprise any of the essentials of economic life' (Pigou 1949: 24–25). Stated in modern terms, this claim has been reformulated as to convey the idea of the

\(^{30}\) See Patinkin and Steiger (1989: 138-141) for an interesting inquiry on the origin of this expression.
neutrality of money on economic activity\textsuperscript{31}: ‘[the property “money is neutral”] characterises the claim that the set of equilibria of an economy is independent of the quantity of money (provided that the latter is always positive)’ (Hahn 1973: 230). As authors like Lucas (1972: 113-114), Niehans (1978: 8), and Patinkin (1987: 639) have put it, even unanticipated variations in the money stock do not affect the equilibrium level of the variables resulting from supply and demand conditions in the real sector.

This issue is however rather controversial in so far as the expected growth rate of the quantity of money is concerned, and it has led neoclassical economists to distinguish between neutrality and superneutrality of money\textsuperscript{32}. In the words of Douglas Fisher, ‘[b]y monetary neutrality we mean the (null) effect of a once-and-for-all change in the quantity of money on the real variables of the economic system. [...] Superneutrality, then, is generally taken to be the (null) effect of a once-and-for-all change in the growth rate of the quantity of money on the real variables of the economic system’ (Fisher 1989: 132). Yet, this distinction seems to contradict the crucial monetarist policy prescription of a constant (possibly low) rate of growth of the money supply, in order to keep inflation on its expected path and hence leave the ‘real’ economy unaffected. On the one hand, what Tobin (1981: 35) labels ideological monetarism promises to rescue our monetary production economies from (unexpected) inflationary pressures, by putting forward the previously mentioned ‘Friedman Rule’ as the specific remedy to the ongoing rise in aggregate prices. On the other hand, by

\textsuperscript{31} Hayek was among the first to claim that money should have a neutral effect on economic activity, so that, following Schumpeter (1954/1994: 277), his analytical picture may be classified as ‘Real Analysis’. In his celebrated lectures at the London School of Economics, published in 1931 under the title Prices and Production, Hayek advocated that the first object of monetary theory ought to be an in-depth investigation of the conditions under which money-using economies behave “as if” they were barter economies. Hence, his policy prescription concerned in principle the appropriate distribution of money balances (and credit) across the economy, rather than the stability of aggregate prices. See e.g. Desai (1981: 33-35) and Patinkin and Steiger (1989: 132-138), but also L. H. White (1999: 110-115, 118).

the very idea of money being superneutral, theoretical monetarism claims that a (unique) change in the rate of money supply growth is harmless for economic performance, for it does not affect any of the real variables in the system. To the best of our knowledge this contradiction has been unnoticed so far and might deserve further investigation. Although the issues involved in it are certainly more complicated than the above antithetical statements, within this paradigm the (super)neutrality of changes in the money supply has generally been treated as being equivalent to a mere change in the unit of account (Haliassos and Tobin 1990: 909; Marty 1994: 407). What exactly a unit of account is will be explored in the next chapter. Here we must consider an even more important argument, which represents perhaps the most striking aspect of established monetary theory, and which is but another way to look critically at the alleged dichotomy.

When he put forth the neutrality-of-money idea in its embryonic form, Professor Pigou pointed out that ‘the number of units of money [...] is, in general, of no significance. It is all one whether the garment, or the veil, is thick or thin. [...] I mean that if, other things being equal, over a series of months or years the stock of money contains successively \( mx_1, mx_2, mx_3 \ldots \) units, it makes no difference what the value of \( m \) is’ (Pigou 1949: 26). Let us refer to the same numerical example made earlier, when we supposed that 100 macro-units of output exist alongside of 2,000 macro-units of money. According to the body of thought spearheaded by Pigou's quotation, the number of these money units may be multiplied by a scalar \( \lambda \) (where \( \lambda \in \mathbb{N}^+ \)) without affecting what he calls the real happenings of economic life. Now, on further thought, one may observe that nothing enables him, or his sympathisers, to maintain that the expenditure of 2,000\( \lambda \) macro-units of money on the product market elicits the power to purchase all the given 100 macro-units of output. So much so that

\[ \text{Say, a decimalisation of the monetary unit (e.g. the changeover from 'old' to 'new' French francs in the 1930s, the British currency reform of the 1970s, the Brazilian decimalisations/renaming of the 1980s, or even more recently the introduction of the single European currency).} \]
the simple juxtaposition of the total stock of money (2,000 macro-units) and the total stock of real goods (100 macro-units) -- two distinct and independent magnitudes in neoclassical analysis -- does not enable one to posit that the purchasing power of the money stock is defined by total output. In this theoretical framework, in fact, 'prices themselves are determined by relating money to product and, therefore, they cannot simultaneously be the necessary condition and the result of this relationship' (Cencini 1988: 171). In other words, in this framework it may be the case that for the purchase of the whole national output a sum greater (less) than either 2,000 or 2,000\lambda macro-units of money is required. By positing a dichotomous economic system, in fact, the maintained method of analysis logically cannot establish the objective value of the existing money stock. "The value of money remains unexplained because it is said to depend on a value [that is, the price level] that can only be determined once exchange has finally related real goods to one another" (pp. 123-124). Conceived in this way, the value of money is (left) indeterminate and, therefore, money cannot be said to be neutral. We simply do not know. As we shall see later, in order to determine the purchasing power of money, and to discover what money neutrality really means, one has to explain how money is associated with national output before transactions occur on the market for produced goods and services.

To sum up, although the modern version of the quantity theory of money has been strongly criticised by different contemporary authors (as will be further examined in Chapter 4), in this section it has been possible to raise two neglected points -- to the best of our knowledge -- which go right to the heart of neoclassical economics. Firstly, monetarists and their most recent heirs, the partisans of cash-in-advance models and of overlapping-generations models, assume output without considering the fundamental problem of how to measure it\(^\text{34}\). Secondly, the same professional

\(^{34}\) Let us stress that this measurement problem has a theoretical, not a statistical, origin. As Sraffa pointed out in connection with the definition of capital and its measure, there exist in fact two distinct types of measurement. 'First, there was the one in which the statisticians were mainly interested. Second there was measurement in theory. The statisticians' measures were only
economists introduce an 'intermediary' good called money, while giving no logical explanation of its purchasing power. Put simply, established models of monetary production economies feature an unmeasured (unmeasurable) output as well as an unvalued (un-valuable) stock of money. Both these fundamental problems come from the dichotomic vision of the economic system, and may be further highlighted by another argument centred on the conventional measure of inflation.

Homogeneity postulate and inflation

Granting the Pigovian (super)neutrality of money, it has been claimed that a change (say, an increase) in the money stock, no matter how distributed within the economy, will produce ultimately -- once all adjustments have run their course -- an equi-proportional variation (increase) in all prices, leaving real variables unchanged (Pigou 1949: 26; Niehans 1978: 2; Patinkin 1987: 639-640). This postulate can be ascribed to Hume (1752/1955: 41-42). It was reformulated by Irving Fisher (1911/1931: 29-32), and it has since been taught in macroeconomics courses by the now classic illustration which asserts that "[d]oubling the money supply doubles all nominal prices but leaves relative prices unaffected, therefore, the volumes of goods demanded and supplied remain constant" (Frisch 1983: 226). Indeed, according to most adherents to neoclassical monetary economics, the fullest sense of the quantity theorem may be stated as follows: 'If, starting from a long-run equilibrium, the stock of money is doubled in the economy, then, regardless of how this money is initially distributed, the economy will not come to rest again until all prices have exactly doubled and every transactor's money holdings have exactly doubled' (Howitt 1974: 143). In

approximate and provided a suitable field for work in solving index number problems. The theoretical measures required absolute precision. Any imperfections in these theoretical measures were not merely upsetting, but knocked down the whole theoretical basis' (Sraffa 1961: 305).

35 This argument might be rephrased by saying that a relative exchange cannot perform the task of measuring the (two) things exchanged, an issue that we shall explore in the next chapter.
abstract formal terms, given two values of the money stock, $M$ and $M^*$, this means that $P$ and $P^*$ are the corresponding (equilibrium) price levels and that $(M^*, P^*) = (\lambda M, \lambda P)$, where $\lambda = M^*/M$ for all $\lambda > 0$. This property has been termed the 'homogeneity postulate' (and it is said to denote the absence of 'money illusion' in agents' economic behaviour), because, in mathematical terms, demand functions are homogeneous of degree zero in money prices$^{36}$.

Clearly, as Morishima observes, the homogeneity postulate is complementary to the quantity theory of money, in all its successive forms and varieties; without this property it would in fact be difficult to envisage neoclassical monetary economics (Morishima 1992: 186). But this also implies that 'once the homogeneity [postulate] is rejected, it is evident that the quantity theory of money does not hold' (p. 187). Applied to our theme, this means that we need to consider carefully what the consequences in terms of inflation are if, say, the money stock is doubled according to the dominant paradigm.

Hence, let us return to our stylised example. Suppose that an imaginary helicopter tosses out another 2,000 macro-units of money over the given national economy, where total output is fixed at the level of 100 macro-units as above. Abstracting from the final distribution of the new money stock among economic agents, how will this hypothetical operation affect the relationship between the stock of money and the stock of output, or, in short, inflation? To illustrate this point, let us write the two situations in parallel, before and after money injection.

---

$^{36}$ This property has also been expressed by saying that the demand function for money depicts a rectangular hyperbola in $(1/P, M)$ space. This means that the product of $M$ and $1/P$ is constant when $P$ (the absolute price level) changes. See Patinkin (1956: 23-31, 39-45) and Niehans (1978: 7-12).
Initial situation

Total output: $Q = 100$ macro-units;
Money stock: $M = 2,000$ macro-units;
Price level: $P = 20$.

End situation

Total output: $Q = 100$ macro-units;
Money stock: $M^* = 4,000$ macro-units;
Price level: $P^* = 40$.

It will not come as a surprise to the reader that the traditional answer to the above question hinges on the variation in the absolute price level. From this analysis one usually infers that the observed increase in the targeted index of aggregate prices represents the underlying measure of the current rate of inflation. Recalling equation (1.1), the actual inflation rate is thus said to be ($\pi$) equal to 100%.

Yet, this view, widely shared as it is, does not contemplate the key problem inherent in the neoclassical dichotomy, as we have been trying to bring forward in this section. Indeed, in this framework, the fact that aggregate prices have doubled does not necessarily imply that the purchasing power of the money stock has been halved. Since the dichotomic account cannot establish the value of the existing money stock (neither does it allow to measure total output objectively), it has in fact nothing to say as to whether or not an increased money supply has an inflationary effect on the national economy. In a situation where the money stock has been augmented, by such bold fabrications as a 'gold rain', mainstream investigation cannot tell us if inflation will result. In such a situation, the only positive information actually conveyed by traditional monetary theory refers to the upward movement in the level of aggregate prices, as estimated by price index analysis, which is rather trivial indeed. This conclusion might surprise the reader, but it is in reality only a first, modest attempt to illustrate what Lord Desai in the "Preface" to his selected essays in *Macroeconomics and Monetary Theory* portrays as 'the greatest obstacle to decent theorizing about a monetary economy' (Desai 1995b: x). To put it clearly, the homogeneity postulate does not belong to modern (monetary) macroeconomics. Neither does the old-
fashioned, quantity-theoretic paradigm based on a dichotomous analysis of our monetary economies of production. To quote Morishima again, 'the theory must be reformulated in a modified, more accurate form' (Morishima 1992: 188). Only after this has been accomplished, along the lines we investigate in the next chapter, can we hope to reach a better understanding (of the causes) of inflation and thus propose the appropriate solution, which a correct analysis of the problem already contains.
CHAPTER 4

THE ARGUMENT REFINED: EXOGENOUS AND ENDOGENOUS MONEY

As we have noted at the outset of this thesis, the characteristic of inflation is to be a condition of generalised excess demand, when 'too much money chases too few goods' according to the monetarists' jargon. The preceding discussion of the fundamental problems of neoclassical inflation analysis may thus be more convincingly elaborated by investigating now the essential origin (i.e. the nature) of money. As a matter of fact, any demand -- excessive or not -- is expressed and exerted in money units. Hence, a rigorous study of the latter can perhaps be useful to cure inflation. To quote from a recent work, '[i]t is a truly well-documented empirical regularity that all persistent inflations are accompanied by a rising stock of nominal money (although recent advances in theory have deprived us of the understanding as to why this should necessarily be so)' (Heymann and Leijonhufvud 1995: 12). To explain this phenomenon, and to determine the underlying mechanism provoking an excessive growth of the number of money units\(^1\), one cannot abstract from both the nature of modern money and its specific role in the economic process. This is a point still worth enquiring into (even after two hundred years of monetary economics), because still '[t]here is no generally accepted definition, among economists, as to what constitutes money' (Vane and Thompson 1979: 49). As Schumpeter imaginatively put it no less than forty years ago, '[t]here is no denying that views on money are as difficult to describe as are shifting clouds' (Schumpeter 1954/1994: 289). This comment is as relevant today as when it was made\(^2\). Too often the basic

\(^1\) In his *New Palgrave* dictionary article, Friedman himself acknowledges that within the analysis of inflation '[t]he deeper question is why excessive monetary growth occurs' (Friedman 1987: 17).

\(^2\) In a similar mood, one may go back to Karl Menger's (1892) *Economic Journal* article "On the Origin of Money", to find out that 'the enigmatic phenomenon of money' is an issue present in the whole history of monetary thought. As Lord Desai elegantly puts it, '[d]ebates over the last two hundred years have used the word money to cover a variety of situations' (Desai 1987: 137). See also Sargent and Wallace (1982: 1212).
idea behind received monetary analysis relies a priori on the functionalist definition of money: ‘money is what money does’ (see e.g. Hicks (1967: Chapters 1-3)). And it need not be recalled here that the conventional representation of the generally-accepted-medium-of-exchange function (the most important quantity-theoretic function) of money merely portrays it as the lubricant that can help alleviate trading frictions, or as the wrapping paper which ‘does not comprise any of the essentials of economic life’ (Pigou 1949: 25). In Hahn’s terms, this raises the challenge of explaining money’s positive exchange value in a framework where money is of no intrinsic worth and is generally accepted in exchange for goods and services (Hahn 1973: 230-236; 1982: Chapter 1). To this question we now turn, in an attempt to illustrate further the congenital defects of the established method of inflation analysis, which is founded, let us stress it, on the dichotomous view spearheaded by adherents to the quantity theory of money (see previous chapter).

EXOGENOUS MONEY, OR THE DICHOTOMY REVISITED

As noted by Kaldor and Trevithick, ‘[o]ne problem which followers of the quantity theory of money had to face from the beginning is the basic question of how “money” is to be defined’ (Kaldor and Trevithick 1981: 11). This section intends to show that this conceptual problem is common to all recent advances within neoclassical monetary economics. So much so that for the purpose of our study we can refer to the exogenous-money view as a kind of portmanteau label unifying a number of established strands of thought. Let us start from the ‘primary definitions’ put forth by the father of the modern quantity theory of money, whose tenets are revived

---

3 As noticed by Hegeland, ‘[t]he very name of Fisher’s well-known formula, the equation of exchange, is a direct indication that money is dealt with only in its function as a medium of exchange’ (Hegeland 1951: 198).

4 Niehans cogently underlines that ‘[e]ach concept of friction [...] implies a corresponding concept of “marketability” or “liquidity” of different goods, measuring their ability to overcome the frictions’ (Niehans 1978: 16).
nowadays by various sorts of money-in-the-utility-function models as well as cash-in-
advance models: ‘Any commodity to be called “money” must be generally acceptable in exchange,
and any commodity generally acceptable in exchange should be called money’ (Fisher 1911/1931: 2). These claims express a common-sense approach. Loosely speaking, money has in fact been considered up to now either as a (perfectly durable) commodity or as an asset, which ‘enables the act of purchase to be separated from the act of sale’ (Friedman 1974: 8) -- a vision stemming from the equation of exchange in its transactions form (see equation (3.2)). This is tantamount to saying that money separates barter trade into two chronologically distinct transactions. It behaves ‘as if’ it were an intermediary good in each of them (see e.g. Clower (1969: 14)). The worrying double coincidence of wants is thus disposed of, as soon as there is something that every agent will accept in exchange as ‘general purchasing power’ and that can serve as ‘a temporary abode of purchasing power’ (Friedman 1974: 9).

Now, the very definition of \( M \) upon which the equation of exchange (and, ultimately, the whole neoclassical monetary theory) is based, appears problematic. Even the most recent formal developments in monetary macroeconomics, such as the great variety of ‘overlapping-generations’ models, seem to be undermined by the same conceptual problem affecting their quantity-theoretic progenitors. Let us try to set out this problem. If the axiomatic starting point of received monetary analysis is a functional distinction between money and non-money goods (that is, the traditional dichotomy in its most elementary sense), then, as Clower pointed out, ‘[we have] to express analytically what is meant when we assert that a certain commodity serves as a medium of exchange’ (Clower 1967: 4).

---

5 It need not be emphasised that every single individual is bound to accept the exchange intermediary in use, for it is known to be acceptable to others by virtue of a tacit social contract. See e.g. Hicks (1967: 5), Clower (1969: 14-15), and Hahn (1982: 21-22).

6 In a different context, Stigum noted the divorce between form and content of mathematical symbols in both economics and economic modelling. ‘To wit: Mathematics is created out of the empty set, and its assertions concern properties of symbols. [...] “[M]oney” and “capital” are symbols just as much as \( x \) and \( y \) are symbols’ (Stigum quoted in Chick 1998: 1861).
The medium of exchange, a monetary asset?

In a 1985 *Journal of Political Economy* article significantly entitled "Money Is What Money Does", a member of the Board of Governors of the Federal Reserve System asserts that '[t]he view of money as means of payment on which the equation of exchange is based provides a relatively clear-cut criterion for which monetary assets to include as components of $M$. In particular, an asset is included if and only if it serves more or less generally as medium of exchange' (Spindt 1985: 180). This statement deserves close attention, inasmuch as it represents the main axiom of any neoclassical monetary model put forth over the last hundred years. Wallace, for one, acknowledges that both money-in-the-utility-function models and cash-in-advance models are 'silent about the qualities an asset must possess in order that it yield utility or serve as a medium of exchange' (Wallace 1988: 35). A conceptual investigation into the fundamentals of mainstream monetary macroeconomics could therefore shed some light on the nature of money, and specifically on its character of general equivalent of goods and services domestically produced. Although these issues cover a well-explored ground, their analysis may be helpful here, for it will serve us as a bridge to some relatively new theoretical insights.

In the most basic terms, according to the traditional postulate, monetary assets are those commodities possessing a means-of-payment attribute, in so far as they are generally accepted in exchange for other goods and services\(^7\). A simple picture might be useful here (Figure 4.1).

\[\text{agent } p \quad \text{good } X \quad \text{agent } s \quad \text{money} \quad \text{agent } p \quad \text{good } Y \quad \text{agent } s\]

\[\text{(purchaser)} \quad \text{(seller)} \quad \text{(purchaser)} \quad \text{(seller)}\]

**Figure 4.1** The medium-of-exchange function of 'monetary assets'

\(^7\) Note that 'medium of exchange' and 'means of payment' are synonymous expressions here (see e.g. the Spindt quote). Every payment is in fact an exchange, as is unanimously agreed within the economics profession, though not on the same analytical grounds. This issue is taken up in the next section.
The representation in Figure 4.1 is so familiar that it does not need any explanation, and the observation that two distinct exchanges are depicted in it might seem trivial. Yet, a number of analytical questions might be raised within this theoretical framework. In particular, where do the ‘initial endowments’ of both money and non-money goods come from? How can prices be formed on the commodity market? More generally, how can the value of the exchanged items be measured? These questions will be considered in the next section. For the time being, the reader’s attention must be drawn to the fact that the exogeneity-of-money approach cannot provide a satisfactory answer.

Let us suppose that money enters the picture like ‘manna from heaven’, i.e. it is dropped from the sky by the high-powered Friedmanian helicopter (Friedman 1969: 4-5). Assume also that the distribution of money across the economy is good enough to feature this (money) asset in each individual’s portfolio. Similarly, assume that every trader has been allocated a set of real goods by, say, the invisible hand, and tries to maximise utility by exchanging (part of) them in separate, or simultaneous, market sessions. Although we accept all this uncritically here, a fundamental problem remains. What is the value of money? How much is a helicopter one-pound note worth?

It is well known that neoclassical (general-equilibrium) monetary economics often relies on the assumption that relative prices — for both money and non-money goods — are known to any agent through a kind of *deus ex machina*, that is, the Walrasian auctioneer. This means that in the whole economy the price system is determined on

---

8 Kocherlakota and Wallace (1998) have recently recast this assumption in the idea that a regularly updated public record of all transactions can inform the optimal allocation of goods and services within a monetary production economy. A closely related paper by Kocherlakota (1998b) aims in fact at formally proving that any allocation which is achievable using an exchange intermediary (the monetary asset) could also be achieved by allowing agents costless access to a historical record (dubbed ‘memory’) of all past transactions. ‘Hence, from a technological point of view, money is equivalent to a primitive form of memory’ (Kocherlakota 1998b: 232). Kocherlakota basically restates the same view when he concludes that ‘money does not reduce the cost of transferring resources from one person to another. There is no immediate technological reason why money
the commodity market, when the relative exchange between money and output takes place (see e.g. Figure 3.1). However, the above fundamental question remains unanswered (not to say it has not even been addressed). In fact, neoclassical theory is unable to explain the value of money, because in this framework the latter 'is said to depend on a value that can only be determined once exchange has finally related real goods to one another' (Cencini 1988: 122-123). If an agent agrees to exchange part of his 'initial endowments' of goods for a number of money units (or for a number of units of the monetary asset), he does so because he knows the value of the received item, which he keeps as a temporary abode of purchasing power, since it will enable him to buy some other real goods in later transactions. 'But how can he assess the precise amount of purchasing power of the received sum if its value depends on the terms with which his commodity is finally exchanged with the other?' (p. 122). Clearly, this question cannot be answered. In neoclassical thinking, value is indeed a relative price. It is determined on the commodity market, where exchange is said to occur between two autonomous objects of trade. Hence, maintaining that the value of money is known when the (relative) value of the exchanged goods has been established, is tantamount to saying that money is inessential in Hahn's (1973) sense. So long as payments are conceived as an exchange between two distinct and independent things -- that is, an exchange involving money and non-money goods in a mutually opposing motion (as in Figure 4.1) --, no theory can explain either the value or the existence of the alleged medium of exchange. How can the value of money be determined, if it is meant to depend on the purchasing power of money, which in the neoclassical 'cosmology' -- to use Verdon's (1996) language -- is established once relative exchanges between real goods and services have ultimately occurred? The

should be a better numeraire than other goods' (p. 250). See also Kocherlakota (1998a) for a less technical presentation of the same argument.
question is clearly circular, as has been pointed out by Cencini (1982, 1988) with respect to the logical indetermination of relative prices.

Further, and closely related to the preceding discussion, one may also wonder how a good, albeit of a transitional or ‘intermediary’ character such as the socially selected monetary asset, can logically be, in one and the same operation, a medium of exchange and one of the objects of trade in every transaction whatsoever. To our knowledge, this point has never been raised before in economic analysis. On reflection, the traditional conception of money as the generally accepted medium of exchange appears ill-founded, and the confusion between means (that is, the instrument, or process, used in performing an action) and end (the object for which the action is performed) is not of a pure semantic order within this context. In fact, as a quick look at Figure 4.1 shows, in the canonical modellisation of a monetary economy money does not play the role of a mere instrumental device -- invented, as a social institution, to help overcome the daunting double coincidence of wants, and to alleviate the practical drawbacks (or frictions) of a non-monetary economy. Money enters every monetary exchange as the counterpart of non-money goods, and, as such, it is an object distinct from real goods and services, with its own ‘mass’ and velocity of circulation, in conformity to the dichotomic view (see Chapter 3). As such, money is not the instrument of payments but the thing which ‘plays a distinctive asymmetric role as one side of virtually all transactions’ (Starr 1980: 263). Clearly, in this framework money is the object of supply and demand as are all non-money goods in

---

9 On less sophisticated grounds, and from a conceptually different vantage point, Verdon (1996: 21-22) raises an analogous criticism. Considering money as a commodity (p. 206, n. 9), he claims that ‘money cannot be both a commodity and a measure of utility because it could not concurrently be something produced and exchanged (a commodity) and a simple medium of exchange and gauge of utility’ (p. 22). Although Verdon seems to point out here the conceptual opposition between objects and means of payments, in his critique of the neoclassical paradigm he does not seem to understand the proper function of the medium of exchange (or of the means of payment). He maintains in fact that ‘there is nothing neoclassical in acknowledging that money is both purchased and [sold], and therefore a commodity, since the money they [neoclassical economists] deal with is only the illusion of a commodity’ (p. 106). See also Verdon (1996: 38-44).
the economy. Indeed, 'money is treated as a stock, not as a flow or a mixture of a flow and a stock' (Friedman 1987: 5). In this approach, money clearly is a store of value, 'a temporary abode of purchasing power' in Friedman's words. This issue is bound to raise with renewed concern the previously noted neoclassical problem of giving money a price, even though, ultimately, it is merely a worthless token used to circulate domestic output. To quote Balasko and Shell at some length on this longstanding puzzle of conventional monetary analysis,

'[m]oney does not in general serve as a proper store of value -- i.e., money cannot have a positive price -- in the finite-horizon economy in which the terminal date is known with certainty. The reason is obvious. Money is worthless at the end of the final period. Consequently, in the next-to-last period, individuals desire to dispose of money holdings in order to avoid capital losses. This drives the price of money to zero at the end of the next-to-last period. And so on. Individuals with foresight drive the price of money to zero in each period, i.e., the 'general price level' in equilibrium must be infinite. The natural way to permit money to be a proper store of value is to go beyond the finite-horizon model (Balasko and Shell 1981: 112-113).

This is perhaps the fundamental reason why so many money demand formalisations (e.g. cash-in-advance models, money-in-the-utility-function models, and money-in-the-production-function models) have been developed over the last decades, along the lines of the portfolio-theoretic approach to asset diversification for finitely-lived 'representative' agents within an infinite-horizon framework.

To take the exploration of the nature of money further in our critical appraisal of neoclassical monetary theory, let us focus on probably the most fashionable form of mainstream economic analysis, namely the overlapping-generations method (OGM). As is well known, OG models investigate the working of contemporary payment systems assuming an intermediary commodity called money, to avert, or at least reduce, search and bargaining costs among the different generations of traders (usually ranged in two complementary classes, labelled 'young' and 'old'). These generations exist in any market period, since they overlap indefinitely as time goes by. On the standard OGM assumptions that (a) before the first-period exchange takes
place (at time \( t \)) all monetary assets are in the old generation's portfolio, and that (\( b \)) the young generation produces, or is allocated, real goods for each market session (see e.g. Lucas (1972: 104-109) and Wallace (1980: 49-60)), and substituting agents \( p \) and \( s \) in Figure 4.1 with respectively the old and the young generation of traders, we get that "[t]he old at date \( t \) could sell their money to the young for commodities, who in turn sell their money when old to the next period's young" (Geanakoplos 1987: 771). Yet, this method of monetary investigation, widely shared as it is, does no justice to the proper idea of the medium of exchange. The two preceding quotations clearly illustrate how the supposed means of payment is modelled as a saleable commodity, that is, as an asset with the highest degree of liquidity, whose existence would be like that of a perfectly durable good, handed down from one generation to the next. In his posthumous *A Market Theory of Money*, Sir John Hicks devoted an entire chapter to "The Nature of Money", where he questioned this vision in the following terms:

> It will no doubt have been taken for granted that in the markets we have been discussing, the typical transaction was an exchange of some article (good or service) for something that was recognized as being money; and it may also have been taken [for granted] that the money was simply handed over, as one does when one buys a newspaper in a shop. A useful way of introducing the monetary theory, which will be the subject of the chapters which follow, is to begin by calling into question these two assumptions, asking how far they are justified (Hicks 1989: 41).

As a matter of fact, following the traditional dichotomy, all neoclassical monetary models assume the use of the selected exchange intermediary to be adequately formalised by the excess-demand function for the money sector, as is the case with the excess-demand functions for goods and services in the real-sector subsystem (see e.g. Patinkin (1956: Chapter 8)). In Fama's words, "[s]ince currency [that is, money] produces real services in allowing some exchanges to be carried out with lower transactions costs, currency has a demand function" (Fama 1980: 50). So, although money is sometimes portrayed as 'the means to the acquisition of a subset of ordinary
consumption goods' (Lucas and Stokey 1983: 82), it nevertheless always enters neoclassical monetary models like a commodity of its own, that is, as a peculiar thing chosen (i.e. demanded) -- for its particular transactions services\(^{10}\) -- within the set of economic objects (goods or assets).

Now, Adam Smith saw it lucidly at the beginning of our science, the means of payment ought not to be considered as an object itself. Although in his time money was reified into a precious metal, which blurred the transcendent distinction between the instrument of payments and its material support (e.g. gold), in his *Wealth of Nations* Smith was already pointing out with admirable clarity the importance to distinguish money proper from money's worth, both analytically and in practice (Smith 1776/1970: 386). In his own words, '[t]he great wheel of circulation is altogether different from the goods which are circulated by means of it. The revenue of the society consists altogether in those goods, and not in the wheel which circulates them' (p. 385). Yet, the idea of money as the instrument of output circulation has unfortunately been lost in more recent economic thought. As a quick example, let us consider the exchange between the whole national output of a given period \((t)\) and the existing money stock. If we adopt, for instance, the overlapping-generations method of analysis, we can apply it either in time or in space, by simply substituting the old and the young generations of traders with, respectively, households and firms. To be sure, any neoclassical model of money-using economies can be represented as follows (Figure 4.2).

![Figure 4.2 The sale of national output in neoclassical analysis](image)

---

\(^{10}\) As a central bank officer puts it, '[i]n addition to general acceptability as means of payment [...] monetary services may consist in many other things, such as liquidity, portability, divisibility, and surety of nominal value' (Spindt 1985: 177, n. 4). See also Laidler (1969: 511).
As is well known, neoclassical analysis portrays the sale (that is, consumption\textsuperscript{11}) of national output as a relative exchange between two dichotomic stocks, money and real goods, which move in opposite directions between economic agents (see also Figure 3.1). The saleable goods (and services, analytically included) leave the firms, where they were stocked, to be appropriated by households. Similarly, the existing money stock is transferred from purchasers to sellers, in a zero-sum process\textsuperscript{12}: 'One man's spending is another man's receipts. One man can reduce his nominal money balances only by persuading someone else to increase his' (Friedman 1987: 4). Stated differently, an offer to sell goods is a demand for money\textsuperscript{13} -- and vice versa --, and both objects of trade are substantial, because they migrate across private endowments. Within such a framework, all that one can analyse is how money and non-money goods end up in individuals' portfolios, after total current output has been sold\textsuperscript{14}. As Ingham forcefully observed, '[t]here is no attempt to account for the "concept" of money as a measure of value (or unit of account) -- or even to recognize that this might constitute an intellectual problem' (Ingham 1996: 515).

So, if one examines in a heuristic manner the concept of excess demand applied to money, one can notice that it is based on a setting where money can be either sold (i.e. supplied) or purchased (demanded), 'as if' it were a good. This is essentially a restatement of the traditional dichotomic representation of money-using economies, which we began investigating in the previous chapter and have been carrying on in this section. It can be rephrased by saying that there may exist a net supply of the $n$th

\textsuperscript{11} Recall that the economic definition of consumption must be distinguished from the physical one.

\textsuperscript{12} To use a famous image, 'money is like the "hot potato" of a children's game: one individual may pass it to another, but the group as a whole cannot get rid of it' (Tobin 1963/1987: 273).

\textsuperscript{13} Since in this framework money is considered as a temporary abode of purchasing power, acceptance of money by a seller of non-money goods is analytically a demand for money (as a demand for an 'intermediary' good).

\textsuperscript{14} Recall that 'invisible goods' such as personal services (e.g. haircuts) can be included in the broader set of real goods, without affecting analysis, because services too must be produced in order to be an object of economic investigation.
commodity, conventionally labelled *numéraire*, which — according to Walras’s Law — is in itself a positive excess demand for the other goods and services taken as a whole. Monetarists consider indeed this phenomenon to be the cause of inflation, so that (general) equilibrium can be restored — in a Marshallian sense — by a rise in the level of aggregate prices only.

However, recalling Smith’s intuition, one could object that ‘the great wheel of circulation’ can logically be neither sold nor purchased; in other words, it is never the object of either supply or demand of its own. In fact, the means of payment must be conceptually distinguished from the things exchanged. To quote a famous opponent of the quantity theory of money, ‘money as such is of no value and does not add to the output of real wealth’ (Hegeland 1951: 241). As testified by the non-addition of money and output in national accounts, the former is not really an element of the set of domestic goods, inasmuch as the value of money and the value of output are not added up to determine a country’s total wealth. In the enigmatic words of De Vroey, which we hope to clarify later on, money is ‘a non-commodity in a universe of commodities’ (De Vroey 1984: 383). Hence, the excess-demand method does not seem to be appropriate for money matters, and the whole analysis of inflation might thus be affected negatively by this approach. To be sure, what Niehans pointed out more than twenty years ago is still valid today: ‘What we need is a theory that treats money not metaphorically *as if* it were a consumer or producer good, but as what it really is, namely a medium of exchange’ (Niehans 1978: 16). This observation is immediately reinforced in a footnote stressing that ‘the real problem is not one of classification but of a better analytical understanding of the functions of a medium of exchange’ (p. 16, n. 39). Thus, on the whole, we are still in quest of the nature of the
means of payment, 'a topic for which monetarists have shown little inclination' (De Vroey 1984: 381)\textsuperscript{15}.

The Cartalist idea of outside money and the inflation tax

In parallel with the postulate identifying the medium of exchange with the 'matter' which carries out this function, it has been alleged that money is essentially an institutional symbol -- a 'creature of law' according to Knapp's State Theory of Money (Knapp 1924: 39-40)\textsuperscript{16}. Within the payment system, money should thus be viewed as exogenously given, because it would exist 'from outside' and independently of the set of real goods, as previously noted. Yet, 'helicopter money' is just a caricatural image of this phenomenon. A less imaginative approach considers money as the only legally valid means of payment (hence the expression of legal tender), for it alone has the imprimatur of the State (see e.g. Wray (1998b)). So, while on one front there are those adhering to the view that the purchasing power of money is somehow linked to its physical representation (labelled Metallists for obvious reasons), on the other front there are those who deny 'the proposition that it is logically essential for money to consist of, say, gold, or to be promptly convertible into gold' (Schumpeter 1954/1994: 288). Advocates of the latter view are labelled Cartalists (following Knapp's terminology\textsuperscript{17}), and their number has increased with the development of modern payment technologies, since the latter are based on money's inconvertibility and make use of intrinsically worthless symbols. In the words of Goodhart, 't[he

\textsuperscript{15} More recently, the same claim has also been made by Ingham, who maintains that 'as an object of study in its own right, it [money] is neglected by the dominant or mainstream traditions not only in modern economics but also in sociology' (Ingham 1996: 508).

\textsuperscript{16} Along Knapp's line, Wray (1998b, 2000) has recently recast this view in connection with Abba Lerner's functional finance approach.

\textsuperscript{17} To be true, Knapp (1924: 32) used the Latin word charta to mean ticket or token, and derived the adjective Chartal to describe the nature of the modern means of payment. 'Money always signifies a Chartal means of payment. Every Chartal means of payment we call money. The definition of money is therefore "a Chartal means of payment"' (Knapp 1924: 38). More recent literature adopts the term Cartalism (dropping the etymological 'h') to indicate this view (see e.g. Goodhart (1997)).
substitution of fiat, paper money, for metallic coin as the main component of currency in the last 200 years provides strong support for the Cartalist view that the monetary essence of currency can rest upon the power of the issuer and not upon the intrinsic value of the object so used’ (Goodhart 1989b: 34).

So, let us imagine an economic system where the medium of exchange is ‘a non-interest-bearing fiat currency produced monopolistically by the government’ (Fama 1980: 50). In this kind of model, money is not convertible into anything else (gold or wheat, for instance), and is never demanded for its own intrinsic value (which is zero, if one abstracts from the very little value of the paper on which it is printed). Several points can be raised.

First of all, at this stage of the analysis it seems almost trivial to emphasise that, contrary to the claim previously quoted, money is not produced, because otherwise it would logically pertain to the category of domestic goods defining national output. When we get right to the heart of the matter, including money among the set of commodities would raise the problem of measuring goods by means of goods, a problem that Ricardo had been trying to solve without success until his death. Whilst it is plain that paper money as such is the result of a production process (i.e. the material result of the printing press) whose costs participate in the definition of national income, the paper is only the physical support (a representative sign) of the medium of exchange proprio sensu. Ontologically speaking, the means of payment is in fact the economic measure (or the monetary form) of domestic goods and services, because it does not have to be measured. In short, it is a unit of account with no dimension whatsoever, as we shall see in the following section.

But in the present discussion the main trouble is probably elsewhere. Provided that ‘it is not legitimate to take fiat money to be an argument of anyone’s utility function or of any engineering production function’ (Wallace 1980: 49), how can it enter, and circulate within, the ‘real’ economy? Fiat money has no intrinsic utility in consumption or production, and its own yield is zero (Haliassos and Tobin 1990:
909). So, where does the value of fiat money come from? Why should economic agents willingly hold it as a temporary abode of purchasing power? Basically, a frequent assumption in modern neoclassical analysis has it that (outside) money is issued by the government -- and distributed to the private sector via banks (hence the use of bank notes) -- in order to collect lump-sum taxes from the individuals (Balasko and Shell 1981: 112, 115-116; Goodhart 1989b: 36; 1997: 17-21; Wray 1998b: Chapter 4)\(^{18}\). Because of the economies of scale involved, the larger the taxes levied on households and firms, the more likely for the State-managed money to enter into private transactions as well. So much so that the 'sovereign power' may impose employing legal tender in such transactions, thus prohibiting the use of any other (private) medium of exchange. Accordingly, the exchange value of an intrinsically useless piece of paper -- 'a ticket that admits the bearer to the great social store of all goods', in the phraseology of Schumpeter (1954/1994: 289) -- would stem from the simple observation that 'one person gives up goods (objects that appear as arguments of utility functions, directly or indirectly) for fiat money only because the person believes that someone else will subsequently give up goods for fiat money at an acceptable rate of exchange' (Wallace 1980: 49)\(^{19}\).

To focus on the issues of immediate interest, one cannot but notice that this view of money is not different from the one investigated in the previous subsection. In spite of the intrinsic uselessness of fiat money, printed on almost costless paper, in this framework money is in fact considered, and modelled, as the counterpart of the real goods and services handed over the counter (Kocherlakota (1998b: 232-233) is a recent example). The neoclassical dichotomy is in the background, by definition, and variations (that is, increases) in the level of aggregate prices are the direct result of an overissue of bank notes, as the traditional interpretation of Ricardo's *The High Price*

\(^{18}\) In some hypothetical environments '[n]o inheritance is possible, so that unspent cash balances revert, at the death of the holder, to the monetary authority' (Lucas 1972: 105).

\(^{19}\) See also Lerner (1947: 313).
of Bullion (1810) has it. Metallists and Cartalists share therefore the same conception of money-using economies on which the established method of monetary investigation is based. There is thus no need to redraw Figure 4.1 here, since no changes are necessary, to depict the exchange of paper money for real goods as neoclassical economists imagine. Hence, let us recall the conclusion already reached earlier on: since fiat money enters standard treatment of monetary production economies like an asset of its own, objectively distinct from (privately) produced goods and services, it cannot logically serve as a means of payment within this framework. It merely represents the historical evolution of the matter on which money is stamped (or marked), and the study of present e-money products along these lines would provide no further insight into the problems investigated here. Since this is an issue lying outside the scope of this dissertation, we shall not pursue it further here²⁰.

However, some important related points should be mentioned at this juncture, for they have often been raised within received inflation analysis. Traditionally, the idea of outside money has been associated with the power of the State to run the printing press, in order to obtain goods and services from the rest of the economy (see e.g. Gurley and Shaw (1960: 68-73), but also Wray (1998b, 2000)). As Keynes maintained at the beginning of Chapter 2 of *A Tract on Monetary Reform*, ‘[a] government can live for a long time, even the German government or the Russian government, by printing paper money. That is to say, it can by this means secure the command over real resources, resources just as real as those obtained by taxation’ (Keynes 1923/1971: 37)²¹. According to this view, the State-issued money enters the economy as the counterpart of domestic output, since it is supplied to the private sector in exchange for real goods and services. Stated slightly differently, this proposition asserts that the monetary authorities never play the role of a *deus ex

---

²⁰ For a modern theoretical approach to e-money, see Piffaretti (1998).

²¹ It should be recalled that in 1923 Keynes still adhered to the quantity theory of money. ‘This theory is fundamental. Its correspondence with fact is not open to question’ (Keynes 1923/1971: 61).
machina in the monetisation of the economy, but exert instead the ‘sovereign power’ sanctioned by law to collect what is known under the term ‘seigniorage’ (see e.g. Black (1987)). ‘Seignorage is a kind of tax, alternative to explicit taxation’ (Haliassos and Tobin 1990: 895). ‘[It] is defined as the amount of real resources that a government acquires in a period simply by virtue of the fact that private agents will hold the currency it prints’ (p. 951). This definition may be inferred from the allegation that the government sector has the monopoly on the emission of fiat (outside) money. It implies that the public sector can undertake expenditure and partly finance the planned budget policies simply by a stroke of the pen (that is, the Treasury printing press). Inflation would thus be the inevitable result of excessive government spending. As Goodhart puts it, ‘[t]he key relationship in the C [standing for Cartalist] team model is the centrality of the link between political sovereignty and fiscal authority on the one hand and money creation, the mint and the central bank, on the other’ (Goodhart 1997: 3-4).

Now, analysis must go further, and deeper, than the mere allegation that ‘[f]iat money is a form of credit where the issuing party is the state’ (Shubik 1987: 317). The inflation tax (or seigniorage) pertains to economic history, when debasement of

---

22 The role of a deus ex machina means that the monetary authorities issue money without asking for any counterpart, a case put by Friedman in terms of ‘helicopter money’.

23 Originally, seigniorage was an amount collected by the seigneur minting the currency out of precious metals. It was equal to the difference between the face value of the currency and the value of its metallic content.

24 In this context it may be noted that many influential economists claim that the central bank is simply an agency of the State. In particular, when the central bank issues bank notes and coins, it would act on behalf of the State: the general government sector would thus acquire a conspicuous amount of resources to finance public works.

25 In recent times, this has indeed been the case on a few exceptional occasions, such as a war or a major natural disaster (e.g. an earthquake).

26 The convergence criteria put forth in the Maastricht Treaty are a recent example of policy guidelines for macroeconomic stability based on this theoretical framework. According to this view, excessive deficits of the general government sector ought to be avoided, in order to reduce the threat of inflationary pressures within the euro area. The required independence of national central banks as well as the compulsory introduction of the ‘no bail out’ clause are clear signs of this approach. See Parguez (1999) for a radical critique of the EMU agenda.
minted coins was a source of funds for the local ruler. A fundamental investigation into the matter should go beyond appearances, to dismiss the yet common belief that seigniorage is still part (albeit a small one) of government revenues. Let us start from the pretended emission of fiat money through the government account in stylised form (Table 4.1).

Table 4.1 The inflation tax: real goods vs. nominal money

<table>
<thead>
<tr>
<th>Liabilities</th>
<th>Assets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Private sector</td>
<td>£</td>
</tr>
<tr>
<td>(fiat money)</td>
<td>Private sector</td>
</tr>
<tr>
<td></td>
<td>(real goods)</td>
</tr>
</tbody>
</table>

As mentioned above, leading professional economists maintain that outside fiat money is the private sector's claim on the issuing government (see e.g. Goodhart (1989b: 287)). Now, since in this view the State-issued money is emitted to acquire real goods (and services) from the private sector, recorded on the assets side of the government's balance sheet, two crucial questions need to be addressed. Does the double entry in Table 4.1 have any meaning at all? In the affirmative, how can we measure the value of the goods (including any financial asset) recorded within the general government sector? Let us try to make sense of these points.

From a quick glance at Table 4.1, one might infer that the result of the operation would be nil, because the same agent, that is, the private sector considered here as a whole, is entered on both sides of the government's balance sheet simultaneously and

---

27 Even in these anachronistic cases, the expression 'inflation tax' is inappropriate and may lead one astray, for the collected amounts (by the seigneur) neither did nor could modify money's purchasing power. Only the distribution of the latter among economic agents was changed by the former, as is always the case when a tax is levied.
for the same amount (e.g. x money units)\(^{28}\). However, as suggested above, it may still be argued that fiat money is the counterpart of real goods. In other words, the State-issued liability -- which circulates as a means of payment among the public and may be returned ultimately to the issuer in payment of taxes -- certifies the deposit of part of domestic output into the Treasury vaults\(^ {29}\). The entry in Table 4.1 would thus not be pointless, for it would record the result of an exchange between equivalents that might be represented as in Figure 4.3.

\[
\begin{array}{cccc}
\text{agent } p & \text{fiat money} & \rightarrow & \text{agent } s \\
\text{(public sector)} & \leftarrow \text{goods and services} & \text{(private sector)}
\end{array}
\]

Figure 4.3  The exchange of fiat money for goods and services

According to this interpretation, the State issues IOUs (i.e. promises to pay to the bearer the amount of, say, x pounds) in exchange for goods and services of the same worth. This view stems however from an oversimplified conception of a monetary economy. It depicts the emission of the alleged means of payment as a purchase for the issuer, an operation that is moreover not synallagmatic, because in the present case the issuing body cannot be sued. Since ‘the recourse of an individual creditor is negligible against the state, but by the law of the state the fiat money must be accepted in payment to extinguish other debts’ (Shubik 1987: 317), this question has been on the agenda of those economists, such as Hayek and the adepts of the Free Banking School, that advocate the ‘denationalisation of money’ as a more efficient, that is, inflation-proof, form of organising a (privatised) payment system (Hayek

\(^{28}\) To simplify Table 4.1, we abstract here from the central bank, which in the present framework is considered merely as an agency of the State.

\(^{29}\) This deposit may be either provisional or definite. In the former case, the private sector will ultimately return fiat money to its issuer, to obtain the corresponding goods provisionally deposited into the Treasury vaults. In the latter case, instead, the State acquires the ownership over the goods deposited -- a legal right epitomised by the famous Latin expression *usu, fructus et abusus.*
Indeed, if our payment systems were built on the foundations depicted in Figure 4.3 (and in Table 4.1), free banking would be a legitimate, though radical, pretension by the private sector economy. In effect, why should the State have the privilege of obtaining goods and services by simply getting its own IOUs circulating within the national economy, that is, without ever really paying at all? Fortunately, the workings of modern payment systems do not abide by this supposed *lex monetae*. As a matter of fact, if the money-issuing mechanism described above were put to practical use, it would be affected by the same fundamental problems as those encountered by the monetarist approach to inflation (see Chapter 3). In particular, were fiat money the dichotomic counterpart of goods and services, no objective measure of the value of current national output would ever exist. As a brief reminder of the previous chapter, consider how commodities newly produced by the private sector could be valued on their own, in economic terms, without resorting to the monetary measure from the very moment of their production. No solution can be provided, because the exchange on the commodity market can only reveal the value resulting from the production process. In a similar vein, one might wonder how the value of a one-pound bank note is to be established in terms of the commodities existing in the natural state, so that the above illustrated exchange between the State and the administered 'populace' would be carried out on an egalitarian basis. It is impossible to answer this question objectively, as long as one does not get rid of the dichotomic view\(^3\).

Of course, we do not deny that the things produced on behalf of the general government sector, that is, public goods and services, can, and in fact do, have an economic measure in terms of money. Yet, this measure exists as a result of the

---

\(^3\) As Wallace recently conceded, 'the challenge we face is to formulate models and policies that are consistent with the essentiality of money and that resemble both actual economies and actual policies' (Wallace 1998: 230).
monetisation of the production process by the banking system\textsuperscript{31}, rather than as a result of an exchange of the kind epitomised by the above depicted swap between (fiat) money and output. Neither on the factor market nor on the product market does such an exchange between national money and domestic output ever occur. To claim the contrary would mean misunderstanding the very concept of the medium of exchange (an issue which we hope to have clarified by now), founding it on the false, dichotomic image of the real world.

Before passing on, and as a gambit, to the next section, which attempts to introduce organically into the economy both entrepreneurs and bankers (who, to borrow an image from Morishima, have been put to sleep in the monetarist edifice), let us try to elucidate briefly the credit granted by the private sector to the State, and to contrast it with the emission of fiat money. Contrary to the above quoted allegation that ‘[it] is a form of credit where the issuing party is the state’ (Shubik 1987: 317), paper money is not generated by law. As Menger would say, ‘[s]anction by the authority of the state is a notion alien to it’ (Menger 1892: 255). Indeed, the State may declare the legal tender by a political decision, but cannot determine its purchasing power, which is the result of economic activity, as we shall see later. This implies that, to understand the nature of money, one has to go beyond both the material and the chartal representations of the means of payment. As the present stage in the evolution of both domestic and cross-border payment systems shows, money is entirely of a banking nature. Current account deposits and bank notes are not fundamentally different, so much so that they can be used indifferently in any monetary transaction (Cencini 1995: 37). When the central bank issues fiat money, it replaces in fact a deposit of the public with private banks (also dubbed deposit banks) with a certificate of deposit in central bank money -- represented by bank notes and/or coins. ‘Taking the place of [commercial] bank deposits, bank notes are one of the

\textsuperscript{31} See next section.
possible representations of a drawing right over national output’ (p. 94), a transformation -- literally, a change of form -- which does not affect its object, that is, the purchasing power of money. By the same token, when the central bank acts as the State's bank (although this is neither its essential nor its specific function), it does so without modifying the existing relationship between national money and domestic output. In real-world economies, seigniorage is an optical illusion. If the general government sector is seeking the necessary funds to finance its current public policies, it must do so in strict compliance with the fundamental rules governing bank money, which stem from double-entry bookkeeping. In particular, all the State can do (and does indeed) to cover the primary deficit of its current account is to issue government bonds on the financial market. As Haliassos and Tobin have it, 'government bonds are simply a means whereby the current generation can undertake expenditure for which future generations will have to pay through increased taxes' (Haliassos and Tobin 1990: 915). As a matter of fact, through the intervention of the central bank, the public sector can advance a purchasing power that will originate in a future production (of either the public or the private sector), by selling bonds to capture current private savings. In practice, this means that Table 4.1 has to be replaced by Table 4.2, which depicts the stylised result of any fund-raising operation carried out by the State on the domestic financial market.

32 As Joan Robinson noticed, '[t]he notes now circulating came into existence as the results of loans from the banks to entrepreneurs, who pay out wages in advance of receiving the proceeds of selling the goods which the workers produce' (Robinson 1956: 226). Bank notes are in fact recorded as a deposit in the central bank's balance sheet (Eichner 1991: 845-846; Lavoie 1992: 164, Table 4.8). There is therefore identity between the stock of money and the sum total of bank deposits existing, at any point in time, in the economy as a whole. The statistical definition of money ($M_0$, $M_1$, $M_2$, $M_3$, $M_4$ etc.) focuses indeed on these stock-magnitudes.

33 The specificity of central banking is to clear outstanding balances in the interbank market and thus guarantee the smooth functioning of national payment systems. See Rossi (1998: 44-50).
Table 4.2 The financing of public sector deficits

<table>
<thead>
<tr>
<th></th>
<th>Commercial banks</th>
<th>Central bank</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>liabilities</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Central bank</td>
<td>£</td>
<td>£</td>
</tr>
<tr>
<td>Private sector</td>
<td>£</td>
<td>£</td>
</tr>
<tr>
<td><strong>assets</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>State</td>
<td>£</td>
<td>£</td>
</tr>
<tr>
<td>Commercial banks</td>
<td>£</td>
<td>£</td>
</tr>
</tbody>
</table>

| **liabilities**                | **assets**      |
| Gov't bonds                    | £                |
| Central bank                   | £                |

Through the sale of (say) Treasury bills, the government acquires the savings formed within the private sector, thus financing its current account imbalance with the intermediation of the national banking system.

All in all, the reality of government finance is fundamentally different from the conventional interpretation put forth by advocates of the exogenous-money view. The private sector buys the newly issued government bonds, that is, it lends purchasing power to the public sector, through an exchange where the possible use of fiat money is just a particular form of transferring ownership over yet unsold output between the two complementary sectors of the national economy. No discrepancy of an inflationary nature can ever exist in such an exchange, because the purchasing power of money is unaltered by the sale of government bonds. Ultimately, what the private sector gives up is the drawing right over a given sum of bank deposits, represented by the number written on paper money (but a cheque would do as well), in exchange for the right to withdraw in the future an equivalent amount of deposits. ‘The check [or the bank note] is merely the evidence of this right and of the transfer of this right from one person to another’ (Fisher 1911/1931: 35). It becomes therefore visible that the idea of outside (fiat) money as the State-issued acknowledgement of debt is ill-founded, because it is grounded on a theoretical framework where the emission of
(paper) money and the financing of deficit-spending units are mixed up in a highly
dangerous way for the socio-political cohesion of any national economy.

We put here aside the theoretical determination of the optimal degree of
government fiscal imbalance, since this issue would carry us too far afield, without
providing any analytical insight into the theory of inflation\textsuperscript{34}. What is needed here, as
Morishima (1992) has argued so convincingly, is to account for the functional
existence of both banks and firms\textsuperscript{35} in real-world economic systems, where the
depression of the purchasing power of money still is a main problem for modern
production economies. Only after this has been accomplished, along the lines we
investigate next, can we probably begin to understand the underlying issues inherent
in inflationary pressures.

**ENDOGENOUS MONEY, OR THE ASSOCIATION OF MONEY AND OUTPUT**

Money, a *numéraire*?

Since the pioneering work of Léon Walras, mainstream economic analysis has been
portraying money as the *numéraire*. To be more precise, two fundamentally distinct,
and opposite, conceptions of the *numéraire* exist in the history of monetary thought,
namely a physical and a numerical conception\textsuperscript{36}. As neoclassical authors model it (see

\textsuperscript{34} One may remember that no particular figure or ratio has ever proved to be grounded on economic
theory, as the controversy over the Maastricht budgetary convergence criteria has well emphasised
recently. See Rossi and Dafflon (1996) on this issue.

\textsuperscript{35} Note that defining banks as a particular 'manufacturing' sector is an improper way to set this
problem. As we are going to see, banks and firms perform two very different, and peculiar,
functions in a monetary production economy.

\textsuperscript{36} Pasinetti (1993) has recently used the distinction between physical *numéraires* and nominal *numéraires* for the analysis of the structural dynamics of prices. Investigating the stability of the
general price level, he namely pointed out that there exists 'an important asymmetry between
monetary regimes in which the numéraire of the price system is physical, and monetary regimes in
which the numéraire of the price system is a purely nominal unit of account, not linked to any
quantitative specification of any particular physical commodity' (Pasinetti 1993: 63-64).
previous section), the *numéraire* is conceived in physical terms, that is, as a commodity (or an asset) taken as standard in an economic system which remains essentially a barter trade system. However, the *numéraire* may also be conceived as a purely numerical entity with no dimension whatsoever, i.e. as a number that measures all the goods and services exchanged in an economy. According to Walras himself, in fact, money as such is a unit of account, that is, a numerical unit used to satisfy the ‘needs of trade’ without being itself an object of trade. In his own terms, ‘the word *franc* [...] is the name of a thing which does not exist’ (Walras 1874/1954: 188). Indeed, the idea of money as a non-commodity has long existed in the history of monetary thought. Professor Pigou claimed that ‘[a] pound sterling is not a thing at all. It is a name handed down in history’ (Pigou 1949: 3). Joan Robinson argued basically along the same lines, when she stated that the pound sterling, ‘in itself, is just a word’ (Robinson 1956: 28). All this conforms to the Smithian intuition that money is essentially a means of payment, or the wheel of output circulation, ‘which, in point of principle, does not add to the amount of riches of a given economy’ (Hegeland 1951: 1).

Yet, despite Walras’s perception that money is nothing but a numerical unit, neoclassical monetary theory has never succeeded in satisfactorily explaining the association of money (numbers) and output. As pointed out by Cencini, ‘[t]he mathematical device of taking a given commodity as standard by making it equal to number one is tantamount to being an illusory trick of arbitrarily associating real output and numbers’ (Cencini 1988: 112). Clearly, ‘*numbers* enter into the [neoclassical] picture only *axiomatically* and for one precise purpose, namely to pave the way for the implementation of mathematical techniques’ (Schmitt 1996a: 105). By applying Walras’s Law to general-equilibrium models of money-using economies, one could indeed always eliminate the money equation from any neoclassical system.

---

37 Notice the implicit reference to Debreu’s (1959) axiomatic analysis lying behind his *Theory of Value.*
of simultaneous exchanges (Cencini 1982: 131-132; 1988: 133-136). In fact, as we have observed in the previous section, 'although Walras does take one of his \( n \) commodities as numéraire (or unit of account) it is an essential part of his theory that the numéraire does not enter into the exchange in any different way from any other of the commodities' (Hicks 1966/1967: 3). So much so that, in Walrasian economics, '[t]he numéraire is not money; it is not even a partial money; it is not even assumed that it is used by the traders themselves as a unit of account. It is not more than a unit of account which the observing economist is using for his own purpose of explaining to himself what the traders are doing' (p. 3)\(^{38}\). In short, within general-equilibrium analysis money is inessential in the sense of Hahn (1973: 231). The medium-of-exchange function of so-called monetary assets is not necessary to determine the models' solution (Rogers 1989: 63). As we have been trying to illustrate above, the alleged medium of exchange is theorised, and modelled, as the general equivalent of non-money goods. It moves, let us say, clockwise in any relative exchange where non-money goods move anticlockwise. To be sure, in such a framework any of the \( n \) commodities will do. As Hicks has it, '[a]ny of the other \( n-1 \) commodities might have been taken as numéraire' (Hicks 1966/1967: 3). Overall, neoclassical monetary models are therefore essentially barter models, since '“money” may always be added [to them] without altering any of the perfect barter results' (Rogers 1989: 46).

Now, the fundamental reason behind the unsuccessful attempts to account for the existence of money by means of received macroeconomic models can perhaps also be explained by the lack of analytical distinction between the functionally different economic agents participating in the production-consumption process. As previously noted, Morishima points to the fact that general-equilibrium theory usually deals with 'representative' consumers only, struggling for utility maximisation under some kind of intertemporal budgetary constraint. It neglects both the firms' and the banks' role

\(^{38}\) See Schmitt (1996a: 110-119) for a development of this criticism.
in the macroeconomy, thus being a mathematical analysis of a dream-world, 'a world which is not totally workable in the context of actual society' (Morishima 1992: 198).

As a matter of fact, '[e]very transaction involves three parties, buyer, seller, and banker' (Hicks 1966/1967: 11). If we analyse the factor market, where the newly produced goods are formed, we can observe that the payment of the 'productive services' involves three poles, that is, firms, workers, and the banking system taken as a whole. The same applies to the product market, because any transaction on produced goods and services requires a seller (usually a firm), a purchaser (usually a household), and a bank as its three constitutive poles. One has in fact always to remember that paper money is just the representation of a bank deposit, and that the transmission of bank notes between agents implies the transfer of the corresponding drawing right (purchasing power) over current production, recorded within the banking system. As stated by Cencini, 'every bank-note corresponds to a book-entry of which it is the mere "image"' (Cencini 1988: 58). And it need not be added that the workings of our sophisticated financial markets obey essentially the same monetary structure, because any payment within the national economy is tautologically a monetary transaction carried out through the bookkeeping records of the domestic banking system.

However, if money and non-money goods were to be exchanged as claimed by adherents to the neoclassical cosmology, the banking system would be nothing but a go-between apt to smooth so-called market frictions. Like the numéraire featuring in any neoclassical model, its role would be inessential ultimately for a money-using economy, since it would originate in a profit-seeking behaviour of a particular type of entrepreneur, dubbed banker\(^{39}\). This is indeed the traditional treatment of money and banking in neoclassical macroeconomics. According for instance to Ball, '[t]he banking system is an intermediary in the sense that it facilitates the transfer of real

\(^{39}\)Hence the conflation of firms and banks into the same functional category in mainstream analysis.
resources from surplus to deficit spending units’ (Ball 1964: 168, italics added). Reference to the ‘banking firm’ would hence serve to distinguish a particular service provider among the set of domestic enterprises. Nothing else. Yet, the crucial question -- generally neglected in monetary analysis -- is to explain how the banking firm, or the banking sector as a whole, can (and does) provide an accounting system of exchange among economic units. This is the essence of modern banking. All the rest is ancillary to that.

The means of payment, a double-entry integer

In his widely-cited Journal of Monetary Economics paper, Fama sets out to explain that ‘the main function of banks in the transactions industry’ is ‘the maintenance of a system of accounts in which transfers of wealth are carried out with bookkeeping entries’ (Fama 1980: 39). He repeatedly observes that, in principle, providing an accounting system of exchange does not require any physical medium or a temporary abode of purchasing power, because the transactions aspect of banking is based on the double-entry mechanism of debits and credits (pp. 39-43)^40. Whilst paper money can be used to represent general purchasing power as well as to transfer its drawing rights among currency holders, the essence of the transactions services provided by the banking system is to rely on purely numerical units -- integers having a concrete economic meaning, as we shall see later on -- to carry out any exchange within the economy. As recognised indeed by Hicks in a 1975 Institute of Economic Affairs Occasional Paper quoted by Laidler and Parkin, ‘money is now a mere counter, which is supplied by the banking system (or by the government through the banking

---

40 Note however both the difficulties and the hesitations of the author -- Professor of Finance at the University of Chicago -- to refuse conventional monetary analysis. In the abstract of his potentially path-breaking contribution, which remains ultimately on neoclassical grounds, he indicates that his paper examines the nature of ‘a pure nominal commodity or unit of account [that] is made to play the role of numéraire in a monetary system [...] and how, through reserve requirements, banks get involved in making it a real economic good’ (Fama 1980: 39).
system) just as it is required' (Hicks quoted in Laidler and Parkin 1975: 742). Though embedded in the improper context of outside money, which we have already dismissed as ill-founded, the idea of (bank) money being nothing but a number issued on demand needs to be further investigated. As a matter of fact, '[t]his number is created by the banking system and this creation takes different forms depending on the technological and institutional context' (Bradley et al. 1996: 112). Since numbers have no economic value as such, the banking system can always issue the exact amount of money units the public is asking for (Cencini 1995: 21). In fact, when providing transactions services to the public, banks make payments as demanded by their clients (hence the endogenous nature of modern money), debiting and crediting them through 'bookkeeping entries [that] are used to allow economic units to exchange one form of wealth for another' (Fama 1980: 43).

Now, the last quotation deserves close attention, for it might represent the crux of monetary analysis. Indeed, taking up Hicks's idea of money as a purely numerical unit -- which might be viewed as a restatement in modern terms of the concept of the 'great wheel of circulation' put forth by Smith --, one may note that book-entry money is not an asset (nor is it a commodity), because no single economic agent would ever give up real goods and services to obtain a sum of numbers (Schmitt 1996b: 97). This is fundamentally the same point raised in the previous section: money is not the general counterpart of non-money goods; it does not exchange with them. On purely conceptual grounds, it makes sense to argue that any exchange concerns a single object, which, as perceived by Fama, is literally transformed by this peculiar operation on the marketplace involving necessarily the issue of the numerical

---

41 J. R. Hicks (1975) "The Permissive Economy", in Crisis '75...?, IEA Occasional Paper Special, no. 43, London: Institute of Economic Affairs (reference taken from Laidler and Parkin (1975)).

42 On a less developed analytical ground, King and Plosser stress in an analogous vein the role of money and banking in reducing the so-called 'shoe leather costs' of individual transactions: 'The financial industry provides accounting services that facilitate the exchange of goods by reducing the amount of time and other resources that otherwise would be devoted to market transactions' (King and Plosser 1984: 365).
'counter' by the banking system. Though difficult to grasp, the analytical importance of what may be called, following Schmitt's terminology, an absolute exchange -- to distinguish it from a relative exchange, where two distinct, and presumably equivalent, objects are traded in the imaginary world of neoclassical economists -- cannot be underestimated. The principle of double-entry bookkeeping imposes in fact that for each single payment the same thing is recorded on both sides of a bank's balance sheet, because what the debtor owes the bank is, ultimately, the object of the bank's debt to its creditor. In other words, the bank's double entry is the numerical measure of the transaction, because the number of money units issued by the bank literally counts the value of the object (ex)changed⁴³.

Let us try to clarify this difficult point by referring to a payment of x money units on the factor market, between a firm and its workers (Table 4.3). In fact, as Lavoie notes, '[m]oney is introduced into the economy through the productive activities of the firms, as these activities generate income. There can be no money without production' (Lavoie 1984: 774). It is thus on the factor market that our analysis must be situated, to understand the essence of money in a production-consumption system (a conceptual step necessary to grasp inflation and its causes).

Table 4.3  The result of an absolute exchange on the factor market

<table>
<thead>
<tr>
<th>Bank</th>
<th>liabilities</th>
<th>assets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Workers</td>
<td>x£</td>
<td>Firm</td>
</tr>
</tbody>
</table>

⁴³ It is worthwhile emphasising that a relative exchange cannot do such a thing, because it is impossible to measure the value of two commodities in a single operation. The principle of equivalence of the exchange-value of two traded objects acquires a heuristic status when allowance is made into the analysis for the concept of absolute exchange. See Schmitt (1966) for development on this topic.
To avoid the temptation to explain a deposit formation by having recourse to a pre-existent deposit (whose origin would remain unexplained), let us make **tabula rasa** of any existing bank deposit. In the period under examination, the payment of the wage bill is carried out without using a pre-existent deposit. It is precisely because such a payment is made from the **tabula rasa** that its result is a net income for the economy as a whole. In fact, as anticipated in Chapter 2, what is gained by workers is an income which is not lost by firms, the latter having no positive income to transfer to the former\(^44\). Now, as Table 4.3 shows, money exists as an asset-liability in any payment, because -- being the numerical form in which both the debit and the credit are recorded -- it appears on each side of a bank's balance sheet simultaneously and necessarily (Schmitt 1975: 13-18). In the words of Wray, '[o]nly if money were dropped from helicopters (or otherwise injected into the system) would it represent an asset without a corresponding liability' (Wray 1990: 13).

This analysis has puzzled, stimulated, infuriated, or annoyed a number of economists (Desai 1984: xi). As we shall stress again later on, the asset-liability nature of bank money does not conform to the statistical definition of modern money used by central banks and international organisations (Chick 2000: 130). At a more fundamental level, it appears to some critics as being 'nothing more than the fact that when a bank creates money it simultaneously creates an asset and a liability' (Watkins 1985: 596). Yet, when one goes beyond appearances, one can notice in this framework the fundamental link between money and production: the substance of the workers' deposit is the newly produced good(s) physically stored with the firm, which is indebted to the banking system for the payment of production costs, as recorded in Table 4.3. As a matter of fact, to make a payment, the bank has to record the object of the transaction both positively and negatively, in one and the same motion, for

\(^{44}\) See Rossi (1998: 28-43) for an analytical elaboration of this point. The case when a firm spends a pre-existent deposit on the factor market will be investigated in Part III.
otherwise the basic rule of double entry would not be satisfied (the whole operation
would be aborted and nullified). In other words, the entry on the assets side testifies
that the firm has a financial debt whose real object is current output in physical terms,
while the entry on the liabilities side records the same good(s) in the monetary form
of a bank deposit owned by the factors of production. Both the debt of the firm and
the deposit of the workers result from the exchange monetised by the bank, 'so that
the deposit created by the bank is balanced by a debt of the same sort' (Cencini 1988:
61). The payment entered in Table 4.3 is thus an absolute exchange, because the same
thing -- a given set of real goods and services -- is the object of both the firm's debt
and the workers' deposit recorded simultaneously in the bank's account. Whilst
workers give up the newly produced commodities in their physical form when the
wage bill is paid, they simultaneously obtain the same items in the numerical form of
(x) money units as their remuneration. This change of form is testified by the
workers' ownership of a claim over a sum of bank deposits resulting from the
monetisation of current production45. So much so that workers can spend on the
commodity market the newly formed bank deposit they own, to obtain the very object
of it, that is, physical values-in-use46. Considering the absolute exchange of current
national output as a whole, and contrasting it with the relative exchange imagined by
neoclassical authors and depicted in Figure 4.2, would thus give the following
representations (Figure 4.4).

45 This can be crudely formalised by writing the identity between domestic output and national income,
that is, the sum total of the bank deposits newly formed -- an issue that we shall explore in the next
chapter.

46 Using the common-sense notion of money, i.e. bank deposits (inclusive of central bank liabilities, as
we have pointed out in note 32), we might illustrate this absolute exchange by an everyday example.
When one gets a cup of coffee, say, by introducing a £1 coin in a vending machine, one changes the
form of the object one owns: before introducing the coin in the machine, one owns the cup of coffee
(or any other equivalent 'real thing', owing to money homogeneity) in its monetary form; after, one
has it in its physical form (value-in-use). Note that we abstract from profits here, since this issue is
not germane to the present discussion. See Part III for inclusion of profits into the analysis.
The absolute exchange on the labour market gives rise to national income, since it associates current physical output with the numerical form issued by the banking system at instant $t$, that is, when the wage bill is paid. Workers are thus paid in an operation that grants them the right to withdraw the result of their own activity in the form of a claim over a sum of bank deposits. Thus then, the latter are not exchanged with the newly produced goods, but are these very goods themselves, moulded into their monetary form (issued by banks -- hence the triangular nature of the payment). In the payment of the wage bill 'workers receive their own product in money. This transaction does not merely define an equivalence but an identity: every worker gets a sum of money which, because of its being issued through the payment of the wage bill, identifies itself with the real output of this same worker' (Schmitt 1984: 347, our
The reverse exchange taking place on the market for produced goods and services shows that until final consumption occurs (say, at instant $t^*$, after $t$) workers -- or, more generally, income holders, if we allow for income redistribution -- hold current output in its monetary form. At $t^*$, when they exert their drawing rights on the product market, income holders obtain in fact, in physical terms, what they previously have been holding in monetary terms: an absolute exchange of the opposite algebraic sign than the previous one leaves them with a collection of values-in-use (i.e. physical things), released from the monetary form in which the latter are subsumed through the banks' instrumental intervention.

So, to focus on the crucial point at issue, one can notice in this framework that money and output are not two distinct, and independent, things (as depicted, for instance, in Figure 1.1), with separate and opposing 'masses' that would require one to determine their contingent equilibrium via the general level of prices. Precisely -- let us repeat it once more, since it will be a central argument for the analysis of inflation developed in Part III --, money as such is the numerical ‘container’ whose freight is given by the newly produced output. In this framework, therefore, money and output move in the same direction, as is always the case for a vehicle transporting its load. Contrary to Clower's axiomatic image that ‘money buys goods and goods buy money’, by which traditional economic theory tries to give money an essential role within the dichotomic vision, money and output are respectively the numerical and the real aspect of one and the same thing, called income. In the remuneration of the production factors, ‘money takes the place of the physical product and becomes its numerical form, so that the exchange between money and output defines their

---

47 '[Dans le paiement des salaires] les travailleurs perçoivent leur propre produit, en monnaie. Il ne s'agit pas simplement d'une équivalence mais d'une identité: chaque travailleur obtient une monnaie qui, du fait de son émission dans les salaires, s'identifie au produit réel de ce même travailleur' (Schmitt 1984: 347). See also Cencini (1988: 83-84).

48 Let us stress again that we abstract here from the firms’ profits. In the last part we will introduce profits into the analysis.

49 The expression ‘vehicular money’ (see Figure 4.4) is meant to recall precisely this effect.
integration: money and output become the two complementary faces of a unique object’ (Cencini 1995: 16).

Now, one might still maintain that a unique object cannot be owned by two distinct agents, namely by a firm and its workers. True. In fact, current output belongs only to its production factors, i.e. workers, from the very moment it is formed by the payment of the wage bill until its possible redistribution through income transfers (e.g. mark-up price mechanisms, direct and indirect taxation) within the economy. Income holders, whether workers or other economic agents who take their place in the banking system’s balance sheet, are the sole owners of yet unsold national output, although the latter is physically stored with firms and recorded within the banking system. Firms owe it precisely because it is the object of their debt to the banking sector (see Table 4.3), and the latter simply matches its credits and its debits so that deposit (i.e. income) holders are the owners of total saleable output at any point in time.

At this juncture, one then has only to recall the clue of present-day monetarism (already investigated in the previous chapter, albeit on different grounds), to show how the maintained method of macroeconomic analysis fails to seize the (endogenous) nature of money and why, consequently, it is unable to tackle the inflation problem in accordance with the fundamental rules of bank money. Let us quote Friedman’s dictionary article once again: ‘Today’s income growth is not closely related to today’s monetary growth; it depends on what has been happening to money in the past. What happens to money today affects what is going to happen to income in the future’ (Friedman 1987: 16). There is no need to repeat here the analysis developed above, to emphasise that the paradigm underlying this view cannot provide an analytical framework for the elaboration of policy guidelines aiming to fight inflation.

50 Let us underline that once output is sold, it does not constitute an object of economic investigation any more (for the sake of argument we abstract here from second-hand sales and purchases). Values-in-use are not objects of economic analysis.
successfully. In so far as money and output are considered, and modelled, as two
different things (and not as, respectively, money proper and money's worth), the
relation between them -- that is, income -- will never appear as distinct and crucial as
it is, both in theory and in practice. Hence, the analysis of inflation as well as its cure
will not be adequate to solve the problem for good. Our next task, therefore, is to
explore an alternative paradigm for inflation analysis.
PART III

A MODERN PARADIGM FOR INFLATION ANALYSIS
CHAPTER 5
WAGE SETTING, CREDIT POLICY, AND INFLATION

In the previous chapter we have noted that viewing money as exogenously given rises
the two intimately related questions of measuring output in economic terms and of
determining the purchasing power of money objectively. These are two important,
and fundamental, problems which the traditional dichotomy between output and
money cannot resolve, as we have been seeking to explain since the second section of
Chapter 3. Now, one might also be interested in investigating whether the supporters
of the endogeneity-of-money approach have been able to fare better in this respect.
Since this question lies outside the scope of this thesis, a constructive answer to it will
have to wait further research. However, to focus on the object of this thesis, it might
be worth investigating here whether the analysis of inflation propounded by the
adherents to the endogenous money view provides a better framework in order to
design and implement an anti-inflation policy that can be successful. Generally
speaking, the endogenous money view has been identified with the Post-Keynesian
School of thought, which is the alternative to the neoclassical code¹ that we have
critically examined in Part II. It is therefore to Post-Keynesian monetary literature
that we shall refer hereafter.

Several authors within the Post-Keynesian tradition consider that inflation has its
source in the difference between the growth rate of money wages and the growth rate
of (average) labour productivity (see e.g. Moore (1979: 57) and Davidson (1988:
167), but also Arestis (1997: 101) and Smithin (1997: 403)). In the light of the
methodological antagonism between Keynesians and monetarists in the history of

¹ 'Post Keynesian economists maintain that the conventional method of analysis, what has been
labelled the neo-Walrasian code, has only limited power to encompass the essential macroeconomic
phenomena of money, uncertainty and time and thus has only limited ability in explaining
macroeconomic problems as high unemployment, persistent inflation and severe financial crises. The
Post Keynesian code is the alternative to the neo-Walrasian code' (Fontana 1999: 174).
economic analysis, what is most surprising is that this widely-shared idea of comparing labour productivity with money wages can be closely linked, as lucidly observed by Moore (1983: 542), to the well-known tenet of the neoclassical (quantity) theory of money. In fact, according to the contemporary incarnation of the latter, to wit, monetarism, inflation occurs, let us recall it, when the quantity of money per unit of output is ‘excessive’ (Friedman 1987: 17). Now, besides the multifaceted conundrum (explored in Part II) of measuring output independently of money\(^2\), which is perhaps the most fundamental expression of the cul-de-sac where the traditional dichotomy imprisons neoclassical monetary investigation, one must also consider how it would be possible to concede wage increases ‘outnumbering’ the wage-earners’ product. To put it in simple terms, is it ontologically possible to measure labour productivity independently of the money wages distributed in remuneration to the workers whose productivity is the object of our measurement? Let us try to investigate this question more closely.

THE MACROECONOMICS OF THE WAGE BARGAIN

Labour productivity, effective demand, and money wages

In The General Theory (hereinafter also GT) Keynes set out to develop a monetary theory of production\(^3\) with no reference whatsoever to either the general price level or its technical alter ego, the aggregate price index. Let us quote at length a crucial passage from Keynes’s most studied book.

The division of Economics between the Theory of Value and Distribution on the one hand and the Theory of Money on the other hand is, I think, a false division. The right dichotomy is, I suggest, between the Theory of the Individual Industry or Firm and of the rewards and the distribution between

---

2 What is a unit of output? How is it to be established in economics? Clearly, one cannot abstract from money when trying to answer these questions (see Chapters 3 and 4).

3 The title of the GT was indeed meant to be A Monetary Theory of Production, a title Keynes chose for his contribution to the 1933 Festschrift für Arthur Spiethoff (Keynes 1933).
different uses of a *given* quantity of resources on the one hand, and the Theory of Output and Employment *as a whole* on the other hand. So long as we limit ourselves to the study of the individual industry or firm on the assumption that the aggregate quantity of employed resources is constant, and, provisionally, that the conditions of other industries or firms are unchanged, it is true that we are not concerned with the significant characteristics of money. But as soon as we pass to the problem of what determines output and employment as a whole, we require the complete theory of a Monetary Economy (Keynes 1936: 293).

It is precisely in order to study the economic system as a whole (its functioning as well as its malfunctioning) that Keynes dealt with "The Choice of Units" in Chapter 4 of his *GT*. As a matter of fact, as noted by Bradford and Harcourt in their recent contribution to *A ‘Second Edition’ o f The General Theory*, the first step required for the elaboration of the complete theory of a monetary economy is to find the appropriate units of measurement for the whole economy (Bradford and Harcourt 1997: 107-109)^4. In particular, Keynes was concerned with finding a measure of production that should be exact and precise for the purposes of a causal analysis (Keynes 1936: 39)^5. In short, since 'the community's output of goods and services is a non-homogeneous complex which cannot be measured' (p. 38) -- because it is an incommensurable collection of miscellaneous things having a multitude of different physical attributes^6 --, Keynes was led, and rightly so, to resort to money as the only means to obtain a homogeneous measure of national output (see Keynes (1936: 41), but also Carabelli (1992: 22-23) and Bradley *et al.* (1996: 124)). Indeed, '[t]he money value of output in a given period is a perfectly precise and determinate

---

^4 See also Wray (1999: 14), who remarks that the choice of units is of fundamental importance to discover essential properties and relations in the macroeconomy. As we shall see later, it is a pity that 'Keynes' proposal on measurement of labour has been ignored by all subsequent macroeconomic schools' (Desai 1995a: 347).

^5 This is the Sraffa (1961: 305) problem mentioned in Chapter 3 (see note 34). Theoretical measures require absolute precision, because any imperfection in them knocks down the whole theoretical basis.

^6 Physical heterogeneity also concerns the physical measure of labour (i.e. the measure of output in terms of labour-time). As we shall see later, Keynes (1936: 41) noted that the physical measure of labour can be made homogeneous by reverting to its monetary measure, i.e. the workers' remuneration. See note 9.
quantity, as money values are at all times strictly homogeneous' (Bradford and Harcourt 1997: 116).

Now, if it is clear that money units are homogeneous when one understands their numerical nature (see previous chapter)^7, one still needs to resist the temptation to make use of index numbers to measure the value of current aggregate output or the purchasing power of money (Carabelli 1992: 24). As a matter of fact, both in the Treatise and in the GT Keynes was well aware that the concept of the general price level is vague and imprecise for the monetary theory of production (see Bradley and Gnos (1991)). So much so that this abstraction cannot be used to deflate the money-value of output, as conventional macroeconomics does with no analytical misgivings.

In Keynes's own provocative words, to be echoed more persuasively by Schumpeter (1954/1994: 1092-1093) two decades later,

the proper place for such things as net real output and the general level of prices lies within the field of historical and statistical description, and their purpose should be to satisfy historical or social curiosity, a purpose for which perfect precision -- such as our causal analysis requires, whether or not our knowledge of the actual values of the relevant quantities is complete or exact -- is neither usual nor necessary (Keynes 1936: 40).

In spite of the standard practice of determining the purchasing power of money by calculating the general price level and its variation over a given time span, this method in fact can only lead to 'mock precision' (p. 40) in the field of macroeconomic investigation, as we have been trying to show in Part I.

---

^7 Carabelli (1992: 23) observes that Keynes does not explain why money is homogeneous. She thus puts forward a tentative explanation based on an analogy between the role of money in the economic process and the role of ordinary language in present-day societies. However, as pointed out by Bradford and Harcourt (1997: 129, n. 16), this analogy is forced and unnecessary. Since money as such is nothing but a number, issued by the banking system as an asset-liability, each unit of money is essentially identical to any other unit of money (existing at the same time), for their origin and nature are the same. £1 today is fundamentally identical to any other £1 existing at the same date, no matter of the support (material or immaterial). The exchange at par on the interbank market is further proof of the uniformity of money units existing at the same time.
Having set himself the task to construct the complete theory of a monetary production economy using neither price index analysis nor the concept of aggregate price level, in the *GT* Keynes proposes 'to make use of only two fundamental units of quantity, namely, quantities of money-value and quantities of employment. The first of these is strictly homogeneous, and the second can be made so' (Keynes 1936: 41). Keynes believes indeed that the economic system as a whole can be positively analysed by focusing on money and labour only (p. 43)*. In particular, the heterogeneity of labour can be overcome by reverting to the monetary measure of its result, in the sense that skilled labour can be weighted in proportion to its remuneration compared to the remuneration of unskilled labour^9. In the words of Keynes, '[w]e shall call the unit in which the quantity of employment is measured the labour-unit; and the money-wage of a labour-unit we shall call the wage-unit.* Thus, if E is the wages (and salaries) bill, W the wage-unit, and N the quantity of employment, E = N*W' (Keynes 1936: 41). This passage is crucial. It might suggest that money and output are associated on the factor market, when the wage bill is paid out by firms to workers with the intermediation of the banking system. In fact, as

---

8 Given the fundamental importance of this question, it is a pity that '[n]o economist -- Keynesian, post-Keynesian, neoclassical or new classical -- has paid any attention to Keynes' discussion in chapter 4 of the *General Theory* as to how to measure employment' (Desai 1995a: 353). As we shall argue, those rare economists who have considered Keynes's choice of units have done so on purely exegetic grounds, that is, within the rather orthodox boundaries of the *GT*. They have thus been trapped (an opportune word) within the confines of a basically neoclassical, exogeneity-of-money view.

9 'For, in so far as different grades and kinds of labour and salaried assistance enjoy a more or less fixed relative remuneration, the quantity of employment can be sufficiently defined for our purpose by taking an hour's employment of ordinary labour as our unit and weighting an hour's employment of special labour in proportion to its remuneration; *i.e.* an hour of special labour remunerated at double ordinary rates will count as two units' (Keynes 1936: 41). On this point, Keynes seems to have gained analytical insight in the transition from the *Treatise* to the *GT*. In fact, when in 1930 he devised the 'earnings standard' as the measure of what he called 'the labour power of money', he noticed that 'it[he] chief obstacle in the way of computing this standard is to be found in the difficulty of finding a common unit in which to compare different kinds of human effort' (Keynes 1930: 63). At that time he thus was probably looking for a kind of average wage index, built ultimately on the same conceptual grounds as those on which aggregate price indices were, and still are, established (see Keynes (1909/1983: 60-61)). See also note 14.

* 'If X stands for any quantity measured in terms of money, it will often be convenient to write X_w for the same quantity measured in terms of the wage-unit' (Keynes 1936: 41, n. 1).
Cencini points out, ‘[w]ithout this association, money would remain empty and output would only be a jumble of heterogeneous physical objects’ (Cencini 1988: 85). This means also that the monetisation of the production process is essentially a triangular operation (see Table 4.3 and Figure 4.4, panel a): it associates the numerical form issued by the banking system and the real goods resulting from the productive services employed by firms. The result of this association is money income (recorded in the form of bank deposits), and it is elicited by the wage bill paid for current production. Thus then, ‘money gives its numerical homogeneity to physical output, and output becomes the content of money income’ (Cencini 1988: 85). This process is the objective measure of economic activity, and all inflationary phenomena must be explained by referring to it, as we shall attempt to show later on.

Yet, before turning attention to this alternative analysis of inflation, a further point should be emphasised, for it can help substantiating the present argument on purely conceptual grounds. As Bradford and Harcourt do, one might question the nature of the wage-unit within the money-cum-labour theory of value put forth in Keynes’s GT. ‘What is the nature of the wage for Keynes?’ (Bradford and Harcourt 1997: 118). If the wage bill results from the association of money and output on the factor market, and considering that labour is the sole true factor of production (Keynes 1936: 213-214), one might still wonder whether wages are the price of labour, as Carabelli (1992: 19, 25) seems to maintain. In the affirmative, a kind of weighted wage index must ultimately be found, in order to account for the great variety of wages paid for different types (or skills) of labour -- a solution arrived at by Keynes himself in A Treatise on Money.10

If so, then would not the attempt to measure average wages be subject to the same pitfalls associated with attempts to measure average prices? If this were the case, then the measurement of the quantity of employment in terms of labour-

---

10 ‘In practice the best we can achieve -- even if we can achieve that -- is to take as our index of the labour power of money or earnings standard the average hourly money earnings of the whole body of workers of every grade’ (Keynes 1930: 63).
units would no longer be precise. [...] In short, Keynes's argument for his choice of units appears to fail (Bradford and Harcourt 1997: 118).

From an exegetical standpoint Keynes's idea of the wage-unit can in fact be interpreted as tantamount to saying that the wage is the price of labour. As a matter of fact, in Chapter 21 of the *GT*, where Keynes deals with "The Theory of Prices", he moves from the case of an individual industry to industry as a whole in the following terms: 'In a single industry its particular price level depends partly on the rate of remuneration of the factors of production which enter into its marginal cost, and partly on the scale of output. There is no reason to modify this conclusion when we pass to industry as a whole' (Keynes 1936: 294). Together with his acceptance of the neoclassical postulate that the wage is equal to the marginal product of labour (p. 5), this might indeed corroborate the interpretation that Keynes did consider the wage as a price (Carabelli 1992: 25; Bradford and Harcourt 1997: 118). According to this view, the wage-unit is essentially a price. So much so that money would be its standard of measure, and we are back to the method used to determine the aggregate price level. Labour, considered here as a particular commodity -- a productive service in the sense of the Marxian labour force --, would be traded against money according to the relative-exchange paradigm on which the neoclassical dichotomy is founded. As Davidson points out, in fact, 'on occasion, Keynes deflates nominal output by the wage unit (the money wage)' (Davidson 1998: 36, n. 5), a method which bears a close relationship with the too familiar one used to determine the purchasing power of money by means of price index analysis. On this basis, it would be quite straightforward to conclude that Keynes conceived the wage as the price of labour.

However, as Bradford and Harcourt mention, theoretically there also exists the possibility that Keynes did not view the wage as the price of labour, although this route has not (yet) been explored by the Post-Keynesian tradition, to which both authors belong. ‘The alternative possibility is that Keynes did not view the wage as a kind of price’ (Bradford and Harcourt 1997: 118). Indeed, a fresh look at the concept
of the wage-unit put forth in the *GT* might indicate that Keynes was proposing an entirely different conceptual approach, which, when further developed, could lead one to a path-breaking conclusion. It is this alternative approach that we wish to explore here. Indeed, the crucial point is not to dwell on what Keynes was really thinking when he put forward the idea of measuring macroeconomic magnitudes in terms of wage-units. This question, quite rhetorical if considered from the vantage point of an exegetic approach, acquires a heuristic status when allowance is made for a theoretical investigation into the working of a monetary production economy. In fact, what counts here is whether an interpretation of Keynes's intuition different from those based on purely exegetical grounds can lead us to a better understanding of the working of modern monetary economies of production.

In simple terms, the idea behind such a proposition is that, moving from the definition of bank money as a mere number, it makes sense to argue that the exchange between money and output on the labour market is a complex operation, namely an absolute exchange which defines their integration in the wage-units thus born as the sole precise and objective measure of economic activity taken as a whole. This line of reasoning has been put forward by Schmitt, who argues that, ‘[u]ltimately, the *number* of wage-units issued is the *measure* of the product of the economy’ (Schmitt 1984: 458, our translation). According to this alternative line of argument, the wage-unit is not simply a unit of money paid out to current production factors. Nor is it a quantity of physical goods produced by any ‘marginal’, or unskilled, labourer. It is both, in the exact sense that the unit of measurement of macroeconomic magnitudes is output moulded into its numerical form by the payment of the wage bill, an operation confirming in practice Keynes's theory of labour as the sole factor of

---

11 As Carabelli observes, ‘[a] complex object needs a complex theory [...]. The problem of the choice of the appropriate units of quantity and of measure [of output as a whole] is the link between these two aspects of theory, that is, object and method’ (Carabelli 1992: 19).

12 ‘Finalement, le *nombre* des unités de salaires émises est la *mesure* du produit de l’économie’ (Schmitt 1984: 458).
production. In short, one cannot be more explicit than this: 'The unit of measurement is the wage unit, because monetary wages define the equivalence of form and substance, that is, of the product and the number of units of money paid out in wages' (Schmitt 1986: 118).

In this alternative framework it is then possible to maintain that each newly produced commodity is measured, as an object of economic analysis, by the number of wage-units recorded in the banking system's balance sheet when banks monetise its production costs. In simple terms, the wage bill paid for the production of any good whatsoever is the economic measure of the latter, since the former is defined by the association of a sum of mere numbers with a 'real thing' produced by labour. For example, good a is worth (say) x pounds and the value of good b is measured by y pounds, because their productive services elicit a wage bill of, respectively, x and y pounds. This may explain why, in the end, '[t]he theory of value and the theory of money are strongly interconnected' (Carabelli 1992: 26). Yet, on further thought, how are relative values determined? Granted that the workers producing (say) good a are paid with x units of money, why is it that the workers producing good b obtain y

---

13 Note that labour is the sole factor of production because only the remuneration of workers elicits wage-units, whose total amount defining the current wage bill is the economic measure of all produced goods and services (inclusive of profit goods). 'So stated, this result seems to be very similar to the classical (pre-classical in Keynes' own terms) theory of value. In fact, it differs from it on a small but very essential point: the monetary measure of labour. According to the classics, labour is measured in physical units, whereas Keynes' basic unit is a monetary one. Wages are then the objective link between money and product, a link that is the direct result of the process of creation called production. Two main difficulties never overcome by the classics are thereby avoided, namely the physical heterogeneity of labour and the integration of money into the real world' (Cencini 1982: 134).

14 That this may be so can also be deduced from a short passage taken from Chapter 10 of the Treatise, where Keynes defines "The Fundamental Equations for the Value of Money". In so doing, he proposes to 'choose our units of quantities of goods in such a way that a unit of each has the same cost of production at the base date; and let O be the total output of goods in terms of these units in a unit of time' (Keynes 1930: 135). Stated more clearly and in the language of the GT, this may amount to saying that the wage-unit is the measure equal to one unit of the production costs of current national output. Indeed, the same claim can also be found at the beginning of Chapter 18 of the GT, where it is stressed that national income is measured in wage-units (Keynes 1936: 245).

15 In other words, the value of each good is determined by the corresponding wage bill. The same is of course true for services (i.e. 'invisible goods').
units of money in remuneration of their activity (with \( x \neq y \))? Once again, the answer can be found in Keynes's work, namely in what he called "The Principle of Effective Demand" (Keynes 1936: Chapter 3). This principle, though expressed by Keynes in a rather complex form, is worth considering, for it lies at the heart of modern macroeconomic analysis\(^\text{16}\).

Let \( Z \) be the aggregate supply price of the output from employing \( N \) men, the relationship between \( Z \) and \( N \) being written \( Z = \phi(N) \), which can be called the **Aggregate Supply Function**. Similarly, let \( D \) be the proceeds which entrepreneurs expect to receive from the employment of \( N \) men, the relationship between \( D \) and \( N \) being written \( D = f(N) \), which can be called the **Aggregate Demand Function**.

[...] The value of \( D \) at the point of the aggregate demand function, where it is intersected by the aggregate supply function, will be called the **effective demand** (Keynes 1936: 25).

Stated differently, this principle means that 'the effective demand is the point on the aggregate demand function which becomes effective because, taken in conjunction with the conditions of supply, it corresponds to the level of employment which maximises the entrepreneur's expectation of profit’ (p. 55). To put it yet another way, 'the effective demand is simply the aggregate income (or proceeds) which the entrepreneurs expect to receive [...] from the amount of current employment which they decide to give’ (p. 55). In short, '[i]t is supply-demand’ (Schmitt 1972: 117), for it is the point where the aggregate supply curve intersects the expected aggregate demand curve, 'in a given situation of technique, resources and factor cost per unit of employment’ (Keynes 1936: 24)\(^\text{17}\).

For the sake of argument let us reconsider a two-sector model such as the one where goods \( a \) and \( b \) make up total current output. Sector I produces consumption

\(^{16}\) Rogers, for instance, notes that 'the principle of effective demand emerges as the key element in a monetary theory of production in the [Schumpeterian] tradition of Monetary Analysis' (Rogers 1989: 178).

\(^{17}\) 'Effective demand is the point of intersection of two curves; it is thus as much a supply as a demand, indifferently one or the other, *identically*’ (Schmitt 1988: 183). Cf. also Chick (1983: 65) and Lavoie (1985: 135-137).
goods \((a)\), and sector II produces investment goods \((b)\). As Keynes has it, ‘[t]he amount of labour \(N\) which the entrepreneurs decide to employ depends on the sum (\(D\)) of two quantities, namely \(D_1\), the amount which the community is expected to spend on consumption, and \(D_2\), the amount which it is expected to devote to new investment. \(D\) is what we have called above the effective demand’ (Keynes 1936: 29). The point at issue is thus the formation and final expenditure of national income as it is created by the productive activity of all firms taken as a whole, that is, in the consumption-goods as well as in the investment-goods sector. Precisely, according to Schmitt (1995-1996, vol. I: 78-79), the principle of effective demand means that sector-II firm pays a wage bill of \(y\) pounds (supply) to its productive services if, and only if, the expected sales of the corresponding output (that is, good \(b\)) -- bought with the newly formed income (demand) -- are equal to \(y\) pounds. If the expected proceeds were higher than \(y\), the firm in fact would not yet maximise its profit by remunerating its productive services by that amount (on the assumption that other firms’ expectations are absolutely fulfilled). As a matter of fact, the sum total of income available in the whole economy would not enable the firm to capture more than \(y\), that is, more than what the latter pays out for the wage bill. Similarly, the same firm would encounter a loss if the expected proceeds from the sale of its output were less than \(y\), i.e. lower than the sum paid to the factors of this production. By the same token, the firm in sector I remunerates its factors of production with a wage bill of \(x\) pounds (supply) if, and only if, it expects a demand for good \(a\) equal to \(x\) pounds. Hence, according to the Schmitt interpretation of the principle of effective demand, the newly formed national income \((x + y)\) will be spent for the purchase of good \(a\) in the proportion of \([x / (x + y)]\) and for the purchase of good \(b\) for the rest (i.e. \([y / (x + y)]\)).
Overall, production is organised according to an effective demand which is merely virtual\(^1\), so much so that firms’ expectations may turn out to be wrong. Indeed, as has been so often emphasised by Paul Davidson, ‘Keynes argued that the economic future was uncertain in the sense that it cannot be either foreknown or statistically predicted by analysing past and current market price signals’ (Davidson 1998: 29). Clearly, what Keynes (1936: 161) called ‘animal spirits’ can influence entrepreneurs’ decisions of production. ‘Once production takes place, however, effective demand becomes truly effective, and its value can no longer be distinguished from that of income’ (Cencini 1984: 177), because it is on that basis that the wage bill is really paid out. The value of output (or the measure of income; see Keynes (1936: 63), but also Gnos (1998)) cannot exist independently of the macroeconomic equivalence that money income establishes between the two opposite flows of its formation (on the factor market) and of its expenditure (on the market for produced goods and services), whereby income is produced and, respectively, consumed through a set of absolute exchanges of the opposite algebraic sign. As noticed by Verdon (1996: 112), the national income received by industry as a whole from the sale of the newly produced goods and services is the very same income which nourishes households’ earnings within the production process. To put it in his own words, ‘entrepreneurs in the short term produce through hiring labour and services and thereby distribute the very incomes from which emanates the demand for their goods’ (Verdon 1996: 112). In a nutshell, effective demand is ‘the amount entrepreneurs will settle to produce’ (p. 112). More precisely, ‘effective demand is defined as a two-way flux, firms “giving” and “receiving” the same product’ (Schmitt 1986: 117). In other words, income formation (i.e. total supply, \(Y\)) and income final expenditure (total demand, \(C\)

\(^1\) See Schmitt (1972) for an analysis of virtual magnitudes and of their importance for income determination.
+ I) are two identical magnitudes\textsuperscript{19}, because otherwise they could not be measured at all in this framework\textsuperscript{20}.

The above argument amounts to saying that any effective demand entails the exact measure of the corresponding output. Any increase in effective demand entails by definition an equally identical increase in output, for, let us repeat, the measure of the latter is precisely given by the wage bill paid out to its factors according to the circular flow of income depicted by actual effective demand\textsuperscript{21}. Ultimately, this interpretation of Keynes's analysis of 'the choice of the units of quantity appropriate to the problems of the economic system as a whole' (Keynes 1936: 37) paves the way for the elaboration of an entirely new approach to the theory of inflation. It suggests, as we shall explore in the next chapter, that inflation might be explained as a macroeconomic disequilibrium between total supply and total demand, in spite of their fundamental identity ($Y = C + I$) as measured in the production-consumption process\textsuperscript{22}.

\textsuperscript{19} This does not mean that full employment is always attained, or that this approach applies exclusively to a Say's Law, real-wage economy (in the traditional interpretation of this law), as stated by Lavoie (1987: 87). Recent works by Schmitt (1984, 1996b) and Cencini (1995, 1996) concentrate on structural issues associated with capital accumulation and income distribution that show how unemployment situations can in reality arise out of a yet imperfect bookkeeping framework used by the banking system to keep track of monetary transactions. More on this later.

\textsuperscript{20} As we shall see later on, some critics of this framework have raised concern with the conception of the period of production adopted in this approach (Deprez 1989: 203), based on the frequency used to pay out wages, as well as with the idea that output is saleable when the wage bill is paid out (Graziani 1989: 139). This framework appears to them as being unable to supply a theoretical explanation of those cases where a good (e.g. a Boeing 777) needs several production periods, defined as above, in order to be manufactured and put on sale on the market for produced goods and services. These cases have been explored by Schmitt (1984: 94-105), who explains how financial claims issued by firms substitute the corresponding fraction of output in any period where output has not yet been completed physically.

\textsuperscript{21} Of course, any physical object (say, a light bulb) produced in period 1 might generate a higher wage bill when produced in period 2, thus conveying the impression of the existence of an inflationary pressure in the latter period. Yet, this is just an illusion. As we shall see in the second section of this chapter, an inflationary disequilibrium can only come into being by an excessive number of money units associated with the same output (that is, output of one and the same period). This is perhaps the most fundamental criticism of traditional inflation analysis, based as it is on the rate of variation of the targeted price index with respect to some reference period.

\textsuperscript{22} It has been argued by Lord Desai that this approach to monetary phenomena 'is still an equilibrium one and leaves one very little room for disequilibrium dynamics' (Desai 1988: xiii). This criticism is linked to the functional isolation of each production period from any other (Deprez 1989: 204), since
Now, by formalising his conception of true inflation in relation to the expression ‘\( MV = D \) where \( M \) is the quantity of money, \( V \) its income-velocity (this definition differing in the minor respects indicated above from the usual definition) and \( D \) the effective demand’ (Keynes 1936: 304), Keynes was led to infer a theory of inflation pretty similar to the neoclassical (monetarist) one. Far away from the potentially path-breaking idea of what in his *Treatise on Money* he labelled ‘profit inflation’ (Keynes 1930: 155), which we shall consider later on, in *The General Theory* he focused on inflation in terms of the stability or instability of (aggregate?) prices as depending ‘on the strength of the upward trend of the wage-unit (or, more precisely, of the cost-unit) compared with the rate of increase in the efficiency of the productive system’ (Keynes 1936: 309).

Several Post-Keynesian scholars have given the inflation problem the same treatment (or a very similar one). Kaldor and Trevithick, for instance, claim that ‘[i]n contrast to the pre-first world war period or the inter-war period, [in the post-war period] average wages have invariably increased faster than average productivity, giving rise to an upward drift in labour costs per unit of output’ (Kaldor and Trevithick 1981: 17). Surely, increases in the physical productivity of labour may be compared to increases in money wages for any given set of wage-earners. However, similar comparisons stem from a yet imperfect consideration of the macroeconomic relationship between money and effective demand. The principle that authors like Kaldor and Trevithick do not consider is that labour productivity cannot but be measured in strictly monetary terms for economic analysis. More picturesquely, labour productivity and money wages are like the volume of a gas and the volume of the room in which the gas is released.

---

in this framework there is no functional link between periods. As a careful reading of Schmitt (1984) will show, however, the approach developed in this part is not dynamic but ‘quantic’ (see Realfonzo (1999: 377)); it is in fact based on quantum time, in contrast with either continuum or discrete-period analysis. See Chapter 7 for more details.
Let us explain this important point by means of a simple numerical example. Consider a manufacture of shirts. Suppose that, owing to technical change, a given worker, W, increases his physical productivity by factor 3: in period 1, before innovation was introduced into the manufacture, W’s labour productivity was (say) 10 shirts per working day on average; after, in period 2, it is 30 shirts per day. Are we going to infer that W’s output is three times what it used to be? An economist’s answer cannot abstract from the worker’s remuneration. If it is plain, in fact, that the number of shirts has been multiplied by 3, it must also be noted that this increase in W’s productivity concerns the physical objects only. What about W’s money wage? To simplify, let us assume that the money wage has not changed from period 1 to period 2, i.e. it is of (say) 100 pounds a day in both periods. With the benefit of Keynes’s insight into the units of economic measurement, we can see that the measure of W’s output has not changed over time either. As pointed out by Schmitt, who refers indeed to Keynes’s choice of units,

[i]t is no doubt true that technical progress enables the national economy to introduce more and more commodities into the form-utility; but output takes on the measure of the form and not the other way round; it follows that technical progress and accumulation of the means of production cannot multiply output in its exact sense: a greater diversity and a greater number of commodities are subsumed under an unchanged output. [...] Output is measured by the wage bill; multiplication and perfection of the means of production are neutral actions as far as the payment of the wage bill is concerned (Schmitt 1984: 495, our translation).\(^2\)

In the case at hand, the wage bill generated in the remuneration of W’s effort is equal to 100 pounds, even after the increase in W’s physical productivity has taken place.

\(^2\) ‘Il est vrai sans doute que le progrès des équipements permet à l’économie nationale d’introduire de plus en plus de biens dans la forme-utilité; mais le produit prend la mesure de la forme et non l’inverse; il s’ensuit que le progrès technique et l’accumulation des moyens de production sont incapables de multiplier les produits en leur sens exact: une plus grande diversité et un plus grand nombre de biens sont compris dans un produit inchangé. […] Le produit est mesuré en unités de salaires; or la multiplication et le perfectionnement des moyens de production sont des actions neutres quant à l’émission des salaires’ (Schmitt 1984: 495).
So, since 30 shirts produced in period 2 are associated with the same number of money units as are 10 shirts in period 1, W's output is measured in economic terms by 100 units of money after as well as before the given innovation is introduced into the manufacture. Strictly speaking, and as astonishing as it might appear, W's output has not been increased over time.24

Consider now the much likely situation where W, arguing from the observed increase in physical productivity, asks the employer for a rise in his money wage. Supposing that this increase is granted, is it (or could it be, depending on the amount) of an inflationary nature? The answer should be straightforward by now. As Cencini states, 'every change in the sum paid to workers simply leads to a change in the way production is measured' (Cencini 1995: 66). Output being measured by the corresponding wage bill, a variation in the amount of the latter is only a variation in the measuring scale of current production (Bradley et al. 1996: 131). For instance, if the wage bargain between W and his employer determines a period-2 money wage of 110 pounds per day for W, the corresponding output in the new situation is measured by this number of money units, no matter of the number of shirts he really produces per unit of effort.25

---

24 Let us refer to a Robinson Crusoe economy in an attempt to convince even the most doubtful reader. Suppose that the production of shirts (by W) represents the whole production of a given, very small, national economy, both in period 1 and in period 2. National income is therefore measured by 100 money units in either period, since the wage bill generated in either period is equal to 100. But so is national output too, because the production of the whole economy has given rise to a total output associated -- through the payment of the wage bill -- with 100 money units (which, needless to say, may be distributed in such a way that, ultimately, the wage-earner, that is, W, can purchase only part of total output, the rest being the firm's profit elicited by the marking-up of retail prices). The fact that the number of physical objects subsumed by the money wages is greater when technological change is enhanced does not affect the fundamental conclusion of our analysis (though it affects, of course, the level of consumers' personal satisfaction).

25 This is not to say, of course, that prices are not going to vary over time. In fact, if in period 1 W's output (10 shirts) was sold at a retail price (x) that possibly covered -- if not exceeded -- factor costs (100 pounds), in period 2 the same worker's output (30 shirts) will probably be sold at a retail price (x', with x' > x) that also covers the increased production costs (110 pounds). Any price index analysis would thus be likely to conclude that an inflationary pressure exists in period 2, since the level of prices could be driven upward (on the ceteris paribus assumption) by the increased price of shirts. Note however that this is not sure yet. In fact, the shirt's production cost has decreased, from 10 to \(3^{1/3}\), pounds, which may induce an analogous decrease in its retail price. As Hicks observed, 'in
The above argument amounts to the proposition that a change in the money wage paid for a same worker's effort (say, an eight-hour day by W) does not modify the relation between the output newly produced and money (i.e. the relation that defines money's purchasing power). Although the shirts produced in the two periods, in the case at hand, might have exactly the same physical characteristics (that is, quality), a period-1 shirt is fundamentally different from a period-2 shirt in strictly economic terms. This is because a period-1 and a period-2 sum of money are not two commensurable magnitudes, even in the case where both sums have same numerical value. So much so that these two shirts could be measured by two different wage bills, as in our previous example, without such a difference being of an inflationary nature.

Yet, no one doubts that the increase in money wages per unit of effort may be ascribed to the workers' attempt at protecting their purchasing power from inflation, whose origin we shall begin to explore in the next section. Indeed, it is a fact that the purchasing power of wage-earners suffers from any loss in the purchasing power of money. Besides the attempt to gain a bigger share of total income in the distribution between real wages and profits, wage increases are actually claimed by trade unions in an attempt to limit the loss in their members' standard of living (in relative and/or in absolute terms). This probably is what led a large number of Post-Keynesian scholars to embrace the so-called conflict inflation approach, a theory to which we now turn in an attempt to point out its inadequacies for the analysis of inflationary pressures. On account of the wide currency of this conflict theory, both in academic quarters and in
public opinion, a critical appraisal of it may in fact be an important step towards understanding the inflation problem and its causes. In particular, the analysis we undertake next may be pretty helpful to underline the fact that observing the evolution of aggregate prices cannot be deemed sufficient to grasp the phenomenon of inflation.

The conflict inflation approach: a critical appraisal

Since Rowthorn’s (1977) influential *Cambridge Journal of Economics* paper on "Conflict, Inflation and Money" was published, an expanding volume of literature attempting to explain inflationary pressures in terms of the conflicting claims on income by different functional classes of agents in the national economy has emerged (see e.g. Kaldor and Trevithick (1981: 16-19), Dalziel (1990), Dawson (1992: Chapter 3), Dutt (1992), Lavoie (1992: Chapter 7), Burdekin and Burkett (1996), Palley (1996: Chapter 11), and the references cited therein). Although the contributions to this literature do not stem all from the same analytical view of the working of a money-using (production) economy and follow different political economy traditions (mainly the neo-Marxian and the Post-Keynesian ones (see Skott (1989)), with fringes such as the French regulationists and, most importantly, the neo-Kaleckians), they all share an important number of essential points as far as the aetiology of inflation is concerned. In particular, they all share the belief that ‘[a]t the heart of the inflationary process is the question of relative income distribution’ (Eichner and Kregel 1975: 1308). This is why, for our present purposes, we can subsume them under the unanimously agreed conflict theory label.

As Burdekin and Burkett pointed out in their theoretical work on *Distributional Conflict and Inflation*, the Conflict Inflation Approach (CIA for short) has to be seen as ‘a methodological approach to conceptualizing, modeling and econometrically testing hypotheses concerning inflation processes in different historical-institutional contexts, not as a particular model of inflation’ (Burdekin and Burkett 1996: 15).
Accordingly, our analysis will centre on the CIA hypothetical framework of inquiry, focusing on the concepts of money, income, and inflation it is based upon. We shall thus not attempt to survey the models built along these lines over the last twenty years or so, since this would be too large a task to undertake here, which moreover would not affect our critical appraisal of the underlying CIA’s fundamental hypotheses.

To concentrate on the latter, it is fair to maintain that the CIA ‘may be viewed as the lineal descendant of the theory of cost-push inflation developed in the 1950s’ (Palley 1996: 182). As a matter of fact, among the possible sources of cost-push inflation suggested by adherents to this theory, by far the best ranked one is that arising from the assumed ‘inconsistent claims on income that emerge from the income distribution struggle between workers and firms’ (p. 182). In this view, the two parties' claims on income may exceed available output at the aggregate level. So, the excess of income claims over national output would be a fundamental causal factor of the observed rise in prices on the market for produced goods and services. Wage-earners try to counteract this upward pressure by bidding for higher wages, thus setting forth an alleged inflationary spiral in which each party seeks to achieve, or to maintain, its targeted income share. Without entering here into either irrelevant details or mere technicalities uninfluential for our conceptual analysis, the CIA encompasses two types of distributional conflicts among the different categories of agents. The first is the well-known conflict between firms and wage-earners, dubbed, respectively, capitalists and labourers by those willing to stress the Marxian ‘class struggle’ in the distribution of national income (see Kalecki (1971)). The second arises within the

---

26 This claim is however a matter of debate among CIA advocates. Dawson states that ‘[t]he conflict theory can be seen as combining demand pull and wage push factors, either one being capable of instigating the inflationary process (or price-wage spiral) but the interaction of both being required for its development’ (Dawson 1992: 26). By contrast, according to the author of *Foundations of Post-Keynesian Economic Analysis* (whose Chapter 7 provides a skilful discussion of conflict theory), ‘[i]nflation is not the result of an objective scarcity [...] . The influence of demand is only an indirect one. [...] Inflation is cost-led’ (Lavoie 1992: 377-378). See also Burdekin and Burkett (1996: 33), and Smithin (1997: 400-401).
working class itself, when each group of wage-earners attempts to re-establish what the latter 'consider to be their rightful place in the social hierarchy' (Lavoie 1992: 414). In short, the two identified types of inflation-generating conflicts are referred to by the price dynamics they set forth, and are thus known respectively under the terms of price-wage spiral and wage-wage spiral. Our aim is to show that they both have the same conceptual problems, which in the end undermine their heuristic status for inflation analysis.

When we get right to the heart of the matter, we can notice that at the core of the CIA lie value judgements about income distribution among functional classes of economic agents. In this approach, in fact, 'inflation is explained by normative values, that is pay norms, customs, equity and justice. These norms have an impact on the perception of what is a fair relative wage, a fair real wage, and a fair profit share. They do have an impact on both the wage-price spiral and the wage-wage spiral' (Lavoie 1992: 379). Indeed, a crucial concept introduced by Rowthorn (1977: 216-217) is the so-called 'aspiration gap'. It indicates that what the parties to the wage bargain think to be just or right -- on account of the set of value judgements referred to earlier on -- does not always correspond to the wage structure actually obtained (see also Hicks (1955: 390), Chowdhury (1983: 649-651), and Lavoie (1992: 378-385) for similar restatements of the same idea). Overall, the share of national income that workers obtain in the wage bargain may be inconsistent with the targeted mark-up set by the firms' pricing policy on the product market. 'It is in this sense that inflation is produced by conflict over income distribution' (Palley 1996: 186). More precisely, as stated by Burdekin and Burkett, 'implicit or explicit conflict over distributive shares is the fundamental basis of a rising price level, which in turn “resolves” the distributional conflict (albeit only artificially and momentarily) via the nominal inflation of the income available' (Burdekin and Burkett 1996: 1). This is tantamount to saying that inflation, here traditionally perceived as an ongoing increase in the general price level, is the escape valve for inconsistent claims over social income
(Dawson 1992: 29). Or, using Davidson's phraseology, '[i]nflation is always and everywhere a symptom of the struggle over the distribution of income' (Davidson 1991: 92).

Similar ideas stem from the Kalecki-Weintraub analysis of income distribution within the national economy (Kalecki 1971; Weintraub 1978). This analytical framework has the merit to put forward a plausible explanation of income distribution in the aggregate, as it has some relevance to a causal analysis of price formation in the retail sectors (where firms' mark-up policies can determine the profit share realised on the product market). However, as Skott notes with cogency, the (Kaleckian) conflict theory of distribution -- which lies behind the CIA -- 'is based on microeconomic reasoning which cannot be extended to the macroeconomic level' (Skott 1989: 40). In fact, a change in the functional distribution of income, as the one resulting from a modified bargaining power between firms and workers, does not modify the amount of national income existing in the period under examination. If it is likely that the real wage and profit shares claimed, respectively, by workers and capitalists are inconsistent, and that current income distribution is the outcome of the relative strength of each class on the factor market as well as on the product market, this 'class struggle' can modify, even marginally, neither total available income nor the money-output relationship established by the monetisation of current production. In other words, it is undeniable that there exists a close relationship between retail prices and income distribution, and that this relation can affect both the level and the 'organic composition' of future production. But this does not imply that a change in the functional distribution of income can lead to the formation of an inflationary gap, that is, a disequilibrium between total demand (i.e. national income) and total supply (national output). 'A change in the distribution of national income only gives rise to

---

27 When monetarism was still in vogue, Davidson used to claim that '[t]he distribution of income is both a cause and a consequence of inflationary processes' (Davidson 1972/1978: 347).
zero-sum transfers: what some agents gain is lost by others' (Bradley and Gnos 1991: 177, our translation).28

To take but a stylised example, let us refer to the situation described by Rowthorn (1977: 219), which may be illustrated as follows. Suppose a well-organised working class, WC, with strong trade unions capable to bargain for very big wage increases.29

We already know that reference to labour productivity cannot provide a theoretical benchmark in order to assess whether or not the wage-push is of an inflationary nature: the product being measured by the wage bill, an increase in the amount of the latter does not alter the relationship between money and output, and therefore it must never be confused with inflation (since in the case at hand no disequilibrium between total demand and total supply can ever be observed in reality). As seen above, the increase in money wages merely is a change in the measurement scale of national output, which identically affects national income at one and the same time. Next suppose, always following Rowthorn's example, that successful oligopolies, SO, dominate the product market. Owing to the SO's position of dominance, capitalists may pursue an aggressive mark-up policy designed to gain a very high share of profit on the market for the goods produced by the WC they employ. Rowthorn's conclusion epitomises the CIA's one, and is worth quoting in its entirety: 'Thus, on the one hand workers are strong in the labour market, whilst on the other capitalists are strong in the product market, and as a result there is a major inconsistency between the two levels of decision-making: workers use their power to obtain big wage increases,

---

28 'Une modification de la répartition du revenu national ne met en œuvre que des mouvements à somme nulle: ce que gagnent les uns est perdu par les autres' (Bradley and Gnos 1991: 177).

29 CIA literature often refers to this case as productivity-led wage increases, although the latter increases may also result from a rapid increase in the firms' profits or in the demand for skilled labour (Hicks 1955: 398-401; Kaldor and Trevithick 1981: 17-18; Lavoie 1992: 415). It is indeed easier for trade unions to ask for wage increases in technological-improving industries. Note also that trade unions cannot ask for a reduction in retail prices, since the latter involve several elements apart from factor cost.
whilst capitalists respond with price increases. The aspiration gap is in consequence very large and there is a high rate of unanticipated inflation’ (Rowthorn 1977: 219)\(^{30}\).

This last conclusion turns out not to be correct, if we define inflation as a macroeconomic disequilibrium between total demand and total supply (and not simply as an increase in the targeted price index). Surely, any decision by firms to rise prices on the product market may shift income distribution in favour of capitalists — *ceteris paribus*. Yet, the decrease in the wage-earners' income share (with respect to the outcome of the wage bargain) cannot be considered as signalling the presence of an inflationary gap, anticipated or not by the various categories of economic agents. As we recalled several times already, total demand exerted by income holders (whether households or firms\(^{31}\)) cannot be modified by income redistribution; neither can the total supply of current national output. Any increase in money wages is in fact a numerical redefinition (via the wage bill) of the wage-earners' output newly produced, and any increase in the selling price of the latter does not affect the relationship between money and output, although it may of course give rise to a higher profit share for the firm(s) (see e.g. Gnos (1998: 44-45)). ‘Defined as that part of income that consumers [wage-earners in the present context] transfer […] to firms, profit does not bite into the purchasing power of money at all’ (Cencini 1995: 52).

Unfortunately, a number of Post-Keynesian authors seem not to consider this important fact yet. They set aside the analysis of money and output -- and of their macroeconomic relation established by effective demand\(^{32}\) --, and seek to explain inflation by the conflict over the functional distribution of income. Dalziel (1990) is explicit on this point: ‘The model [he proposes] is stripped of all unnecessary detail (for example, […] money) to highlight the underlying mechanism that allows inflation

\(^{30}\) According to Rowthorn (1977: 215), were inflation (correctly) anticipated by all concerned agents, it would have no redistributive effect, for then everyone would take advance measures to allow for future price increases.

\(^{31}\) For the sake of brevity we abstract here from the general government sector.

\(^{32}\) Chick's monetary analysis is a noteworthy exception. See e.g. Chick (1982, 2000).
to resolve income distribution conflict' (Dalziel 1990: 426). Palley (1996) is another
typical example. In line with Sidney Weintraub (1978), he ‘suggests that incomes
policy, which is designed to reconcile the conflicting claims of workers and firms, is a
superior policy [in respect of traditional monetary and fiscal policy measures designed
to curb inflation]' (Palley 1996: 199).

Now, the incomes policy called forth by the CIA stresses incidentally that
monetarism not only is an anachronistic theoretical construct; it is also bound to lead
domestic policy makers into a dead end, because controlling the money supply cannot
limit agents' claims for a higher share of national income. From the CIA’s standpoint,
the more the money supply is restrained (to try to counter demand-side inflationary
pressures), the more the individuals' income claims on available output may be
inconsistent in the aggregate. Both in relative and in absolute terms, the struggle for
each unit of income would in fact become sharper the fewer money units are in
existence. Yet, by emphasising functional income distribution and especially
conflicting claims on total saleable output, the Post-Keynesian approach seems to be
grounded ultimately on the same fundamental, quantity-theoretic hypothesis on which
monetarists have built their theory. More precisely, the theory of conflict inflation
embodies the neoclassical dichotomy between the stock of money and the stock of
output -- as overtly embraced by Rowthorn (1977: 235-236), who starts in fact from
the basic quantity equation when discussing the role of money within the CIA
(p. 230). Indeed, it cannot be denied that in any CIA-based model 'the inflationary
gap is still fundamentally considered as an excess in the quantity of money over the
physical volume of goods and services' (Bradley et al. 1996: 128). In this framework,
using the CIA's own language, inflation involves an excess of money-income claims
on available output, and 'must be conceptualized as the difference between aggregate
real income claims (expressed in monetary terms as the total nominal value of
contracts comprising the income side of GNP) and real income' (Burdekin and Burkett
On the whole, one may therefore conclude that all those theoretical approaches seeking to explain inflation by social conflicts over the distribution of income among functionally distinct (classes of) economic agents suffer from the same conceptual problems of their monetarist alter egos (see Part II). The alternative offered by the CIA’s theoretical framework lies indeed between (1) a theory of inflation à la Dalziel (1990), where money plays such a passive role that it is (or can be) abstracted from when measuring macroeconomic magnitudes (like total output), and (2) a theory of inflation à la Burdekin and Burkett (1996), where money plays a more active role because it can exist independently of national income (see e.g. Rowthorn (1977: 229-230)). The first point of view suffers from totally neglecting the essential role banks play in a monetary production economy. The second theoretical standpoint relies instead on a partial -- if not superficial -- understanding of this role, which to the best of our knowledge has never been dealt with (not even in passing) in the voluminous conflict inflation literature. Both urge on a deeper money and banking theory to inform the analysis of inflation. As we shall see in the next section, the bank credit mechanism could have in fact a part to play in the explanation of the formation of an inflationary gap, that is, a disequilibrium between total demand and total supply of current output.

Before turning our attention to this important topic, on which much research is needed (particularly at a time when speculation activities in financial markets represent an increasing share of total transactions), let us attempt to consider the problem of income distribution and inflation within the theoretical framework used in this thesis. It will be seen that an increase in retail prices induced by inflation (on the assumption that an inflationary disequilibrium has been created by some yet unexplained mechanism) shifts income distribution in favour of capitalists, since it

---

33 Morishima (1992) and Realfonzo (1998) analyse several examples of this neglect in twentieth-century economic theory.
34 See Howells (1999), and Howells and Hussein (1999).
increases the profit share. It will also be seen that no price-wage spiral can actually be considered as a cause of inflation, because both price and wage increases are implemented in order to limit the loss in money's purchasing power (due to this yet mysterious force dubbed inflation) and are therefore a consequence -- not a cause -- of the latter loss. The conclusion will be a restatement of the fact that price-wage spirals can in no case modify the relationship between money and output.

Let us consider a very stylised macroeconomic model and assume that the class of workers as a whole, WC, faces the class of capitalists as a whole, CC. Suppose that, as a result of the wage bargain, WC is remunerated with 120 macro-units of money for the total production of the current period. We already know that output as a whole is measured by the total wage bill paid out by firms for its production. So, the 120 macro-units of money have the necessary and sufficient power to purchase total current output. Two points need to be addressed at this stage, the first with respect to the conflict approach briefly investigated above, and the second in connection with the problem of inflation that we shall explore in the remainder of this work.

Let us imagine that CC seeks to obtain 25 percent of total income as monetary profit of the period. To this end, CC sets retail prices on the product market higher than the corresponding factor costs, marking them up by the targeted profit rate (25%). For instance, an item whose production cost is 20 pounds has a selling price of 25 pounds. Overall, on the assumption that each and every good is put on sale at a retail price 25 percent higher than its factor cost, the expenditure of the whole amount of nominal wages (recorded in the form of a sum of bank deposits) enables WC to obtain only 75 percent of total current output, since to buy the latter altogether an amount equal to 150 macro-units of money would have been necessary (in the case at hand no other deposits exist apart from those formed in the remuneration of the current production factors). If we analyse WC's expenditure of 120 on the product market, we can observe that 75 percent of it covers the production costs of output.
sold, the remaining 25 percent being a transfer of income in favour of CC. By marking-up the selling prices of current output, CC is able to capture a monetary profit (24 macro-units of money) whose purchasing power is exactly defined by the stock of yet unsold output. On the whole, the income created on the factor market (120) is partly destroyed on the market for produced goods and services (96) and the rest is captured by firms as their monetary profit (24).

Now, suppose that, owing to inflation, there is a general increase in prices (ultimately, this is the symptom-based definition of inflation). Applied to the stylised example investigated above, this means that output is sold at a higher price than in the previous case. Assume therefore that the very same output as above (worth 120, since 120 is the sum total of the wage bill paid for its production) is now put on sale at a price of 180 (instead of 150). If we abstract again from any other income that might exist in the whole national economy apart from the bank deposits corresponding to WC's remuneration (120), we can observe that the expenditure of the latter on the product market enables WC to obtain a lower share of total output than before. To wit, WC gives up a sum of bank deposits equal to 120 to obtain in exchange goods and services whose production costs amount to 80 only. In fact, the share of real wages in respect of total income (120) is reduced by the inflationary rise in retail prices, to the benefit of the profit share captured by CC on the market for produced goods and services. In the case at hand, 40 units out of the 120 spent by WC are captured by CC as monetary profit, the rest (80) covering production costs of sold output. Although this result is identical to the one arrived at by conflict theorists in terms of functional income distribution, it differs from the latters' result in so far as

---

35 Since in this numerical example prices are 25 percent higher than production costs, a part of total output worth 96 (as measured by the corresponding wage bill) is sold at a price of 120. By selling it at a marked-up price, CC is therefore able to cover the relevant production costs (96) and to gain a profit (24) on top of that.

36 See next chapter for an investigation of the link between income distribution and capital accumulation.
the causal link between income distribution and inflation is concerned. Conflict theorists would probably argue that the struggle for the distribution of national income elicits an inflationary rise in prices (setting forth a price-wage spiral), because the functionally distinct categories of agents cannot agree on the relative shares of total output. The approach adopted in this thesis suggests instead that an inflationary rise in prices (whose origin has yet to be explained) shifts the current functional distribution of income in favour of capitalists, CC, thus reversing the causality between income distribution and inflation put to the fore by the CIA. It need not be underlined here that this reversed causality also challenges the symptom-based definition of inflation introduced in Chapter 1, and adopted by conflict theorists, because at the analytical level it is no more sufficient to define inflation by its effect on prices.

It is particularly in an attempt to study an alternative paradigm for inflation analysis that we shall devote the remainder of this dissertation to investigating the monetary macroeconomics of inflationary disequilibria. Given the importance of this problem, we hope that our attempt will arouse the interest of the academic as well as the policy maker, and prompt to further research in monetary analysis for an inflation-proof economic system.

**BANK CREDIT AND INFLATION**

As some very rare ‘Monetary Keynesians’\(^3\) such as Victoria Chick, Hyman Minsky and Paul Davidson have pointed out, the credit facilities provided by commercial banks to the economy may be granted beyond the level that would ensure monetary equilibrium. In other words, the relationship between money and output might be affected by an excess of bank credit, which would thus contribute to the inflationary

---

\(^3\) We borrow this felicitous expression from Desai (1989: 172; 1992: 120). It echoes the Monetary Analysis approach that Schumpeter (1954/1994: 276-288) used to oppose to Real Analysis. A brilliant exposition of both approaches can be found in Rogers (1989).
rise in goods and assets prices in actual markets. Linking the conflict inflation approach to the real bills doctrine\(^{38}\), Davidson (1988: 166-168) maintains that the banking system can theoretically provide support to, if not generate, inconsistent claims on available output. More precisely, when he states that wage increases might be granted beyond changes in (average) labour productivity, Davidson holds the banking sector as co-responsible for the issue of so-called ‘inflation bills’. As he puts it, ‘[a]ny healthy banking system apparatus which meets the needs of trade can be subverted to create an elastic currency of “inflation bills” rather than “real bills”’ (Davidson 1988: 167). So much so that, in the view of Burdekin and Burkett, ‘Post Keynesian models of the endogeneity of the stock and velocity of credit-money with respect to firms’ money wage bills are really particular specifications of the monetary accommodation of excess income claims’ (Burdekin and Burkett 1996: 33)\(^{39}\).

The monetary accommodation (by banks) of excess, or inconsistent, income claims is thus considered an inborn feature of endogenous-money systems, and must therefore be pursued further here to assess its macroeconomic character in respect of inflationary pressures.

Excess credit facilities and inflation: a reconsideration

In the face of ‘endogenously changing financial practices’ (Palley 1996: 219), microeconomic theories of bank behaviour have been continuously revised, and improved, to account for those phenomena related to the bank intermediation

\(^{38}\) The real bills doctrine has been traced back to Adam Smith’s *Wealth of Nations*, where one can read that ‘a real bill of exchange [is] drawn by a real creditor upon a real debtor, and [...] as soon as it becomes due, is really paid by that debtor’ (Smith 1776/1970: 402). As Green notes, ‘echoes of the real bills doctrine reverberate in modern monetary theory’ (Green 1987: 101). This is essentially because money emission is traditionally understood as if it were an exchange between a bank’s IOU and a non-bank ‘bill’, which may be backed by ‘titles to real value or value in the process of creation’ (p. 101), or it might not. See e.g. Sargent and Wallace (1982) and de Boyer (1998: 65-67, 74-75).

\(^{39}\) See also Burdekin and Burkett (1996: 14).
As a matter of fact, since the rapid development of liability management, banks in general have been inclined to lend following a policy of aggressive expansion of their balance sheets, each time the new credit facility is deemed profitable for their business strategy (Chick 1986: 115-118). Recently, in a series of papers, Howells (1995, 1996, 1997, 1999) has depicted this state of the art very clearly. He namely points out with cogency that what Davidson describes as ‘needs of trade’ (see above) are not the exclusive causal factor of bank lending. ‘Certainly the idea that new bank lending originates solely with firms and reflects their production plans now seems very naive’ (Howells 1996: 113). The relative increase in speculation, and hence in total spending, with respect to the value of current output is indeed a reality of modern capitalist economies which can no longer be denied. Howells’s conclusion ‘is that bank lending results in a monetary expansion that may be quickly accommodated by multiplier-assisted increases in income, or it may not’ (Howells 1995: 96). In simple terms, if the new loans granted by banks are not associated with goods and services newly produced, the question Howells forcefully raises is to determine whether or not the ‘deposits created by essentially speculative activity have

40 See Screpanti (1997) for a recent attempt to work out a structural theory of endogenous money from a microeconomic point of view, where ‘the core business of bank activity is identified in the transformation of generic risk’ (p. 567).

41 Broadly speaking, liability management refers to the ability of banks to increase their lending activity by borrowing funds which appear on the liability side of their balance sheet, without having to dispose of their marketable assets -- mainly Treasury bills’ (Lavoie 1992: 212). See Moore (1989: 13-20) and Goodhart (1989a: 30-32) for a discussion on the importance of liability management in the evolution of modern banking.

42 See also Howells and Hussein (1998, 1999).

43 As Keynes already observed (partly quoted in Howells (1999: 105)), speculative transactions need not be, and are not, governed by the volume of current output. The pace at which a circle of financiers, speculators and investors hand round one to another particular pieces of wealth, or title to such, which they are neither producing nor consuming but merely exchanging, bears no definite relation to the rate of current production. The volume of such transactions is subject to very wide and incalculable fluctuations, easily double at one time what it is at another, depending on such factors as the state of speculative sentiment; and, whilst it is possibly stimulated by the activity and depressed by the inactivity of production, its fluctuations are quite different in degree from those of production’ (Keynes 1930: 47-48). See also Chick (1994).
any impact on the economy, different from the impact of deposits created in the wake of production’ (Howells 1996: 113).

In order to explore carefully the macroeconomic significance of this concern, let us analyse a stylised example of excess credit. As in Howells’s work, let us focus on agents’ borrowing to purchase second-hand housing assets, although any kind of speculative (that is, purely financial) demand for loans will do here. For the sake of exposition we shall assume that a single deposit bank represents the whole network of commercial banks, and that the bank-accommodated demand for loans exceeds by 10 percent the deposits corresponding to, and resulting from, the monetisation of current production. There are no pre-existent deposits. So, if a worker (or any income holder, IH, if we allow for income redistribution within the economy) asks for, and obtains, a bank loan in order to purchase a second-hand housing asset whose price exceeds the worker’s (or the income holder’s) deposit, in bookkeeping terms the situation may be represented as follows (Table 5.1).

<table>
<thead>
<tr>
<th>Table 5.1</th>
<th>A speculation-led excess of bank credit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bank</td>
<td>liabilities</td>
</tr>
<tr>
<td>(1)</td>
<td>Worker (IH)</td>
</tr>
<tr>
<td></td>
<td>Firm</td>
</tr>
<tr>
<td>(2)</td>
<td>Seller</td>
</tr>
<tr>
<td></td>
<td>Worker (IH)</td>
</tr>
<tr>
<td></td>
<td>Borrower</td>
</tr>
</tbody>
</table>

As explained in Chapter 4 (see e.g. Table 4.3), the first double entry (1) is the result of the monetisation of the worker’s current output, which occurs at the level of, say, 100 money units according to Keynes’s principle of effective demand illustrated earlier on. Current income holders, IH, thus have a purchasing power of 100, saved
in the form of a bank deposit until consumption takes place. Now, if the bank were to grant a credit of 10 to current income holders (for instance, to our worker) for, say, a house purchase (see e.g. Howells (1997: 431-432)), entry (2) would be recorded in the bank's bookkeeping as soon as the purchase is made. Yet, maintaining that for any bank loan there is a corresponding deposit is only saying the obvious fact that any double entry affects identically, at one and the same time, the assets side and the liabilities side of a balance sheet. As a result of the payment, in fact, the 10 units of money borrowed by the bank's client are instantaneously deposited by the seller of the house, who will have to choose the form in which he prefers to hold his wealth (expressed by a sum total of deposits equal to 110). If he decides to buy existing and/or newly issued financial assets, the corresponding deposit is transferred to some other unspecified agents, who will then face a similar choice. Be that as it may, the 'last holder' of the deposit will spend it on the market for produced goods and services, so that, in our stylised example, 110 units of money will be used to purchase the very same output the production of which gave rise to a (wage-earners') deposit of 100 -- in this example, no other output is available for purchase.

Assuming, to simplify what does not affect analysis, that total current output (worth 100) is sold at a price of 110 to the 'last holder' of the deposit which results from the credit granted to the borrower, the corresponding transactions are recorded as follows (Table 5.2).

---

44 Recall that when consumption occurs, income holders spend their deposits and thus extinguish the firm's debt, which was recorded either towards the bank or with respect to its own wages fund (working capital). In either case the firm is able in fact to cover on the product market the previously disbursed factor costs.

45 As Keynes (1930: 42) remarked in *A Treatise on Money*, one ought to bear in mind that the amount of a customer's unused overdraft does not appear anywhere at all in a bank's statement of its assets and liabilities.

46 It is worth while to emphasise here that this subjective choice is necessarily subsequent to the deposit automatically formed at the very instant the payment is made.
Table 5.2  The macroeconomic result of excess credit facilities

<table>
<thead>
<tr>
<th>Bank</th>
<th>liabilities</th>
<th>assets</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>Worker (IH)</td>
<td>100</td>
</tr>
<tr>
<td>(2)</td>
<td>Seller</td>
<td>110</td>
</tr>
<tr>
<td>(3)</td>
<td>Client C</td>
<td>110</td>
</tr>
<tr>
<td>(4)</td>
<td>Firm</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td>Firm</td>
<td>10</td>
</tr>
</tbody>
</table>

Entry (3) records the purely financial transaction between the seller of the house and another client, C, of the bank (the ‘households buy from households’ case in the words of Howells (1997: 431)). It depicts the fact that the seller’s liquidity preference leads him to buy some kind of financial assets from another household. Entry (4) records instead the result of the ‘households buy from firms’ case (p. 431), and epitomises the consumption of current output (sold at a mark-up over its production costs)\(^\text{47}\). The overall result is that the firm makes a profit (10) because of the presence of a net borrower within the economy; this might lead one to infer that an inflationary pressure has been generated by the bank’s excessive lending. In other words, there is an excess of aggregate demand, caused by too much money chasing too few goods. The observed increase in output prices (and hence the firm’s ‘windfall’ profit) should be sufficient proof of the likely consequences of the ‘passing around’ of ‘extra-money’, using De Vroey’s (1984: 384-389) language which we shall consider later.

\(^{47}\) Note that in both entries (3) and (4) client C might also be a firm, selling financial claims in order to purchase other firms’ output. Within the time-dimensional sequence depicted by entries (2) and (3), we would therefore experience the case where ‘households buy from households and lend to firms’ (Howells 1997: 431).
on. To quote Howells again, 'u
less everyone has an overdraft, unwanted deposits may continue to circulate. It is precisely this that gives rise to those repercussions on prices, quantities, of goods, assets or whatever' (Howells 1995: 94).

The worries raised by Howells are legitimate, and certainly deserve careful attention by all those concerned with income distribution among economic agents. We indeed already noted in passing that the functional distribution of income has a close link with economic growth, too. We may even claim that the latter is probably path-dependent on the former. No-one could in fact deny today that the investment of (purely speculative) profits in (perhaps off-shore) financial markets is a less productive activity, at least in terms of employment, than what has become conventional to label 'physical' investment. Surely, purely financial transactions can generate no additional income, but only transfer existing income among the parties (both in space and time). Only production can give rise to a net income within the national economy as a whole. This thesis, however, does not focus on economic growth. Nor does it focus on the distribution of income. It aims in fact at investigating the origin and cure of inflation, as well as at evaluating (as a by-product) recent monetary research, theory, and policy.

So, to focus on inflation, we have to concentrate attention on 'the inflationary potential of a banking system on which there appear to be few constraints' (Chick 1986: 122). As pointed out by Chick, '[t]he banks' aggressive lending activity may contribute to inflation' (p. 123), because the balance between assets and liabilities in the banks' account is not sufficient to prevent excessive credit to be granted.

48 In this connection, as every economist knows, several authors have studied the relationship between inflation and growth. See e.g. the survey paper by Driffill et al. (1990), and Barro (1996).

49 A similar point has been raised by De Vroey in connection with unsuccessful business ventures, when he maintained that the endogeneity of money is 'not a sufficient guarantee of the eventual correctness of private money creation decisions' (De Vroey 1984: 388). See also Smithin (1997: 397-398), who emphasises the role of bank credit for consumption purposes (beyond the need for the firm's initial finance). On the concept of 'initial finance', and its link to 'final finance', see Lavoie (1987: 69) and Graziani (1990: 14-16), who elaborate on Keynes's (1937) Economic Journal finance motive. See also Chick (1996: 14) for the importance to distinguish 'finance' from 'funding'.
Moreover, the widespread use of liability management and the pursuance of credit policies inspired by this business strategy have further enhanced the lending power of commercial banks, which have no innate structural device to prevent the financial bias towards inflation (p. 124). All this, indeed, should be worrying enough to consider attentively the nature of credit-induced deposits with only a remote link to production. In short, is this superfluous of bank credit definitively inflationary for the national economy? Is the firm’s profit (equal to 10 in the above example) irremediably inflationary, because it is due to an over-emission of money? Are income holders alienated, because their deposits are allegedly spoiled by the banks’ sales- or profit-maximisation policies beyond a kind of notional threshold determined by the income-generating finance process à la Davidson (1988)?

Certainly these concerns are justified, in so far as they highlight the fact that the relationship between money and current output could be altered by excess bank credit. However, from a fundamentally macroeconomic point of view, using the Schmitt terminology their object may be assigned the label of ‘benign inflation’, in the precise sense that this kind of alteration of the money-output relationship is not irreparably inflationary. In other words, the inflationary effect of excess bank lending is not cumulative in time. Let us attempt to explain why, referring to the stylised example made above.

When a borrower is granted credit, he actually engages himself to repay the loan back at the due date (inclusive of any interest payment agreed upon). The bank’s recording of this loan in the assets side of its balance sheet testifies that the borrower relinquishes simultaneously a financial claim, perhaps by simply acknowledging his debt, i.e. by endorsing the loan agreement in front of the banker. Now, what is the final object of the claim disposed of by the borrower and recorded within the bank’s

---

50 See also Minsky (1975: 121-124).
assets? This claim is a title to a future income that the borrower will earn and pay back into the lending bank. Indeed, as Cencini argues, 'the working of our banking systems is such that overdrafts of private banks are reduced to an advance' (Cencini 1995: 62). As such, the monetisation by banks of borrowers' claims, as in entry (2) in Tables 5.1 and 5.2, does not irremediably modify the relationship between money and output. The increase in existing bank deposits entailed by the advance of a borrower's future income will be matched, in fact, by a correspondingly identical reduction of money income when the borrower reimburses the bank.

The importance of this argument is likely to impel us to prove our conclusion further, or more firmly. Consider the consequence of the loan repayment for the macroeconomy. When the excess of bank credit (which we know to be an advance of future income) over the value of current real goods falls due (say, in period \( n \)), the borrower has to surrender an equivalent part of his period-\( n \) income. Clearly, if the loan granted in period 1 must be repaid in period 2, a period-2 income has to be transferred from the (period-1) borrower to the bank. And when the reimbursement of an outstanding bank loan occurs, an equivalent deposit is simultaneously extinguished; a destruction also noticed by Howells, when he claims that 'only actions that cause repayment of loans cause a reduction in deposits' (Howells 1995: 100). Now, there is no real need to illustrate this destruction in bookkeeping terms, to infer that the reduction in (for instance) period-2 deposits amounts to an identical decrease in total demand. In fact, since any demand can only be exerted through an

---

52 This does not imply that the borrower is the original holder of the income used to reimburse the bank. It might well be the case that this income has to be borrowed from period-\( n \) income holders. In this case, the borrower once again anticipates a future (period \( n+m \)) income, so that there simply is a substitution of due dates, leaving the overall situation unaffected. In what follows, we assume therefore that the borrower can repay his debt in period \( n \) with no further commitment, since the opposite hypothesis is not germane to the present analysis.

53 This is the link that money provides between the present and the future, as noted by Keynes (1936: 293). More recently, Minsky (1994: 154-155) has reasserted the link that financial relations provide between the past, the present and the future, enabling deficit-spending units to consume today what they will earn in the future.
expenditure of bank deposits\textsuperscript{54}, destruction of the latter \textit{ipso facto} reduces current demand by the same amount. So, whilst in period-1-like cases total demand may be increased by excess bank credit such as to create an inflationary gap, in period-2-like situations this gap is compensated by a disequilibrium of the opposite algebraic sign, total demand falling short of the level established by the monetisation of current (that is, period-2) production. 'Hence, even if we claimed that private bank overdrafts are a cause of inflation since they modify the relationship between money and current output, we would have to add that the discrepancy between demand and supply which it causes is not seriously worrying, for it is bound to be compensated for in the following periods' (Cencini 1995: 63).

This conclusion may be strengthened by three different arguments, each of them stressing the benign character of 'credit-led' inflation. Firstly, '[a]t any particular time, existing loans are being repaid while new loans are being demanded' (Howells 1995: 90). Present-day monetary production economies are indeed so complex that in any period of time new credit facilities are provided simultaneously with the repayment of outstanding loans\textsuperscript{55}, 'so that the positive gap between demand and supply caused by bank overdrafts is normally balanced in each single period by an equivalent gap of the opposite sign' (Cencini 1995: 63). The very large number of commercial banks operating within modern domestic economies is further guarantee that no serious inflationary threat can come from the bank credit mechanism: whilst some banks might grant an excess of credit in respect of the same banks' deposits, some others may lie behind or even encounter excess reserves\textsuperscript{56}. Secondly, '[m]ost of the time the excess of credit granted by a bank leads to it becoming indebted to

\begin{flushleft}
\footnotesize
\textsuperscript{54} Recall that bank notes and coins are the material representation of a bank deposit, to wit, a deposit in central bank money.

\textsuperscript{55} See Arestis and Howells (1999: 118).

\textsuperscript{56} Theoretically, if all deposit banks acted together to grant, overall, excessive credit, then inflation would exist until the latter excess is absorbed by the borrowers' reimbursements. However, as we shall see in the remainder of this subsection, this highly hypothetical case is very unrealistic in modern banking.
\end{flushleft}
another bank, a situation which banks try to avoid as much as they can' (Cencini 1996: 29). The ongoing multilateral clearing of interbank debts in the central bank's accounts is indeed a very efficient, and highly developed, mechanism aimed at minimising any excess of credit a single bank may grant over any particular period of time. Referring the reader interested in a macroeconomic analysis of modern interbank settlement systems to the work of Cencini (1995: 31-46), let us very concisely note here that even the alleged lender-of-last-resort function taken on by the central bank has no inflationary bias when properly understood (Rossi 1998: 44-50). Inasmuch as commercial banks foster the distribution of income among the non-bank public without modifying the money-output relation, the central bank has no way to act but as a mere go-between in the interbank market. Being analogous to chemical catalyses, central bank's interventions within the national banking system neither do nor can affect domestic monetary equilibrium. Carried out on behalf of the reporting deposit banks, these interventions never really lead to inflation, because at worst they serve to advance a future income which will be destroyed when the corresponding output is formed. Thirdly, advances of income via the banking system (that is, commercial banks and the central bank taken as a whole) are beneficial to economic growth. They accelerate in fact the sale of current output, enabling firms to cover their production costs (either by restoring their working capital or by repaying bank loans), and thus prompting an expansion of activity. The positive influence of this mechanism on the level of employment, and growth, is rather straightforward to infer in this case, so that the identifier 'benign inflation' should be definitively justified to characterise the topic of this subsection. To put it clearly, 'speculative' bank lending may be a cause of inflation since it affects the relationship between current output and money. But any inflationary gap elicited by an excess of bank credit is destined to be absorbed by a disequilibrium of the opposite algebraic sign when these excessive loans
are reimbursed, that is, when bank deposits are reduced by loan repayment\(^5\). In a nutshell, benign inflation is not cumulative in time.

Yet, the reader attentive to the rhetoric of the argument may eventually wonder if there also exists a kind of 'malign inflation', whose wording sounds already much more threatening than its altogether beneficial compeer. We need therefore to turn attention to this problem, and the remainder of this thesis should be regarded as a modest attempt to dig out a fundamentally unexplored ground in inflation analysis (monetarist as well as Keynesian).

Inflation: is there any macroeconomic harm?

So far, excessive money emission has been associated with the bank credit process supporting agents’ exchanges of goods, services, and financial assets. Each time a bank grants a credit \textit{ex nihilo}, that is not based on past or actual production (as in the real bills doctrine), it might put monetary equilibrium at risk, inasmuch as the extra amount of deposits thus formed adds to the stock necessary and sufficient to circulate current output. Yet, as we have seen, the inflationary bias of this phenomenon is not cumulative in time, and is quite restrained in space also. In fact, credit-led inflation calls forth its own absorption within the macroeconomy. So much so that focusing on the evolution of assets and commodity (aggregate) prices over time can provide no relevant insight into the analysis of the pathological, irremediable alteration of the money-output relationship.

To develop the latter analysis, therefore, neither ‘the state of trade’ (\(P_y\)) nor the amount of total monetary transactions (\(P_T\)) can be a fruitful object of investigation. Expressions \(P_y\) and \(P_T\) subsume indeed exchanges occurring between items which

\(^5\)Needless to say, this does not exclude that in reality a new inflationary gap might arise out of the new loans granted by the banking system at precisely the same time when the previously granted loans are reimbursed. It is this continuum of newly granted bank loans that makes it ‘empirically’ impossible to observe the absorption of benign inflation.
exist simultaneously in the personal endowments of the contracting parties\textsuperscript{58}. Since any exchange is identically a sale and a purchase for both parties, no disequilibrium between supply and demand can logically be conceived of, either at the disaggregate or at the aggregate level, at any point in time. In so far as attention is focused on exchange, the very concept of excess demand that underlies the inflationary gap is therefore going to lie outside the explanatory power of economic theories. As we have observed earlier, in fact, the distribution of income -- hence of output --, both in space and time, can have no inflationary bias at all in the whole economy. This means that the conflict view of, and the accommodative approach to, (potentially) inflationary phenomena need to be replaced, ultimately, by a macro-theoretical analysis of the process by which money income is formed. In a nutshell, to discover the origin of macroeconomic disequilibria in the money-output relationship, production rather than distribution ought to be the focus of analysis.

The above argument amounts to the proposition that there is no analytical bridge between either \( Py \) or \( PT \), on one side, and \( Y \), on the other side (see Schmitt (1972: 15-56) for a logical demonstration that income is not a price). Both \( Py \) and \( PT \) are of a microeconomic nature and are directly related to an exchange economy. They imply aggregating prices into a particular price index, \( P \) (see Part I for a critique of this method). As such, they include transfers of income or profits earned by entrepreneurs at the expense of consumers (Keynes 1930: 301). And it need not be recalled here that no aggregation of microeconomic magnitudes can be interpreted as a macroeconomic outcome. Per contra, \( Y \) is a macroeconomic magnitude from its very formation. It pertains to a monetary economy of production. More precisely, as noted earlier, it is the measure in terms of the wage bill of the output newly produced (see e.g. Gnos (1998)).

\textsuperscript{58}In this theoretical framework, an illustration of this phenomenon is the exchange between total output and the existing money stock on the market for produced goods and services. See Chapters 3 and 4 for a fundamental critique of this view.
Now, whilst it is impossible to find a gap between total demand and total supply as they can be observed in the whole national economy (recall the previous exploration of Keynes's principle of effective demand), it is nevertheless possible to conceive of a macroeconomic disequilibrium within the action whose result is the fundamental identity of income and output. This claim might appear quite extraordinary at first. We need therefore to focus attentively on the very process by means of which income (and output) is formed. In other words, we have to investigate the realm of flows, as opposed to the analysis of stocks changing hands within the national economy, before taking up the inflation problem again. Indeed, to repeat it, the object of a macroeconomic analysis of inflationary pressures must be production, not circulation, for those interested in discovering where the origin of these pressures lies.

When we get right to the core of the analysis, we observe that the formation of income, that is, the creation of the relation between money and current output in the banks' bookkeeping, is essentially an exchange of flows. In the previous chapter we already noted that payments on the factor market are absolute exchanges: workers abandon their product in kind, to receive it in the monetary form issued by banks. Thus, money intervenes within the flow of production, and not only when output has already been formed and is exchanged on the goods market. As pointed out by Schmitt, '[m]oney creation and production of commodities are one and the same action' (Schmitt 1984: 450, our translation). In fact, the measure of output as a whole (which is given by the total wage bill of the relevant period, let us emphasise it) is established by the substitution of the real flow of goods and services with the monetary flow of nominal wages, at the precise instant when the wage bill is paid. So,

---

59 By contrast, output circulation is an exchange of (pre-existent) stocks (see also previous note).
60 This is essentially a restatement of the most fundamental critique of the quantity-theoretic approach to monetary production economies, as discussed in Part II.
61 'La création de monnaie et la production de biens sont une seule et même action' (Schmitt 1984: 450). In English, "each monetary flow and the corresponding real flow are the two components of one and the same "motion"" (Schmitt 1996b: 96).
‘instead of producing real goods directly in their physical form, the economy produces real goods in the monetary form’ (pp. 458-459, our translation).

Now, provided we distinguish money from income, only a small analytical step is necessary to draw a first important conclusion (which in the previous chapter had been left implicit on purpose). Income, or the power to purchase the output newly produced, is the result of the exchange taking place on the factor market between the monetary flow (money) and the corresponding real flow (production). It is because the payment of the wage bill does not necessitate a pre-existent income that it has a positive-sum result for the economy as a whole. The income newly generated is a macroeconomic magnitude in the exact meaning of the term, for the income earned by workers is lost by nobody else. It is true that firms obtain the newly produced output when paying out the current wage bill. But they match their increase in stocks (+) with an identically equivalent increase in their financial indebtedness (−), either to their own wages fund or to the banks. Ultimately, their gain is zero on the labour market. Workers, on the other hand, are the exclusive owners of total current output at the very instant the wage bill is paid, and in fact their gain is positive for the entire ‘working class’ as well as for the whole economy. In the language of Schmitt, ‘[r]emunerations of factors of production are negative expenditures both for the firms and their employees’ (Schmitt 1996a: 137). In other words, negative expenditures are income-generating for the national economy as a whole.

According to this line of thought -- which some authors have traced back to Keynes -- expenditures on the market for produced goods and services are then objectively positive for all economic agents. This implies that, ‘irrespective of the observer, such expenditures are [...] income destroying [...] for the national economy.

---


63 In fact, the firms’ profit is realised on the market for produced goods, as we have seen in the discussion of the conflict inflation approach (see previous section).

64 See Cencini (1988: Chapter 4) and Gnös (1992a: Part III).
and for the whole world' (p. 137). Therefore, considered from the point of view of society as a whole, the link between money and output established by the remuneration of production factors is definitively severed by income expenditure on the product market. In particular, in so far as the deposits formed in the payment of the wage bill are expended for the purchase of the corresponding output, firms can cover their factor costs by balancing the debt originated in the labour market with the deposit obtained on the market for produced goods (see e.g. Cencini (1988: Chapter 6)). Income formation (that is, production) and income destruction (consumption) are taken indeed in a macroeconomic circuit by the circular flow depicted by effective demand, as we have already noted in the previous chapter.

Yet, at this stage, the reader may be somewhat puzzled, if not irritated, by the exclusion of profits from the preceding analysis. This exclusion is only apparent. Our previous investigation of income distribution and inflation should have provided a case where profits are explained by the circular flow of nominal wages between firms and workers, linking the factor and product markets. The macroeconomic analysis developed by the Schmitt School shows in fact that profits are subsumed under the production-consumption process described earlier in terms of the circuit of income (see e.g. Sadigh (1988), and Gnos and Schmitt (1990)). It may be useful to recall here that profits are an income that wage-earners, or, more generally, income holders transfer to firms on the product market (via the mark-up price mechanism). More precisely, the amount of bank deposits transferred from the buyer to the seller of newly produced goods and services that exceeds the corresponding factor costs is not destroyed in the economy. It persists in the banks' accounts and represents the firms' profit. Hence, firms always have the purchasing power necessary and sufficient to acquire the whole physical output defining the real content (i.e. the substance) of their monetary profit.

Now, granted that income redistribution on the product market cannot alter the relationship between money and output, that is, monetary equilibrium, one has to
consider what are the likely consequences if firms do not spend their profit on the market for produced goods and services. In fact, firms may well decide not to consume their profit, that is, they may want to invest it instead of distributing it to their shareholders. If so, then firms can invest their profit either (1) on the financial market or (2) in the production process. When the firms' profit is invested on the financial market, 'at the end of the chain' there will always be a seller of financial assets who consumes his earnings in the purchase of the output corresponding to the firms' profit. This case is therefore not interesting here, for logically no discrepancy between total supply and total demand can thus be detected in the whole economy. To be true, from a macroeconomic point of view the firms' investment of profit on the financial market is not an investment at all. It is consumption, since the substance of the firms' monetary profit is finally consumed by an unidentified seller of financial assets. This case is therefore similar, ultimately, to the consumption of redistributed profits by the firms' shareholders and entitled agents. We are thus left with the investment of profit on the factor market, an issue whose macroeconomic implications are far from simple to investigate and to deal with, as we shall see.

Let us try to set out the conceptual framework of our analysis, to be developed later.

It is well known that a firm's payment of current factor costs does not necessarily give rise to a proportional increase in the firm's debt to the banking sector. It can and is in fact much likely to be nourished (in part) with a pre-existent profit, that the firm invests on the labour market for a new production. 'But then the payment of wages becomes a twofold operation: monetisation of new production on one side, and expenditure of pre-existent income on the other' (Cencini 1995: 73-74). Such a twofold operation can be represented in terms of the macro-objective distinction between negative and positive expenditures described above (Figure 5.1).
As shown in Figure 5.1, an income-generating expenditure may include an income-destroying expenditure. For instance, a payment of a wage bill equal to 100 pounds could be financed up to 10 percent (say) by accumulated profits and 90 percent by new bank loans.

Now, to focus on the issue of immediate interest, it is important to observe that the expenditure of a firm's profit on the factor market gives rise to an anomalous result. This anomalous result does not arise because of the immediate consequences of this operation on capital accumulation (which may be altogether beneficial to growth and employment). It arises because the actual recording of this expenditure in bank accounting leads to the emission of what one may define as 'extra-money' (De Vroey 1984) or, more properly, as 'empty money' (Schmitt 1984; Cencini 1995, 1996, 1999). To the extent that it pays the current wage bill out of pre-existent income, the firm in fact purchases its workers' activity (hence its result also), 'so that workers are effectively credited with a sum of "empty" money' (Cencini 1996: 53), that is, with an amount of deposits to which no new saleable output corresponds.

Let us consider a simple numerical example to try to illustrate the problem further. Suppose that in period 1 a firm makes a profit equal to 10 pounds, say by marking-up its products' retail prices. Were the firm to redistribute this profit (to its share or equity holders, or even to the general government sector in payment of taxes), the latter would ultimately be used up for the final consumption of period-1 yet unsold output. The fact that there exists an unsold output matching the newly formed profit

---

**Figure 5.1 Inclusion of a positive expenditure into a negative expenditure**

As shown in Figure 5.1, an income-generating expenditure may include an income-destroying expenditure. For instance, a payment of a wage bill equal to 100 pounds could be financed up to 10 percent (say) by accumulated profits and 90 percent by new bank loans.

Now, to focus on the issue of immediate interest, it is important to observe that the expenditure of a firm's profit on the factor market gives rise to an anomalous result. This anomalous result does not arise because of the immediate consequences of this operation on capital accumulation (which may be altogether beneficial to growth and employment). It arises because the actual recording of this expenditure in bank accounting leads to the emission of what one may define as 'extra-money' (De Vroey 1984) or, more properly, as 'empty money' (Schmitt 1984; Cencini 1995, 1996, 1999). To the extent that it pays the current wage bill out of pre-existent income, the firm in fact purchases its workers' activity (hence its result also), 'so that workers are effectively credited with a sum of "empty" money' (Cencini 1996: 53), that is, with an amount of deposits to which no new saleable output corresponds.

Let us consider a simple numerical example to try to illustrate the problem further. Suppose that in period 1 a firm makes a profit equal to 10 pounds, say by marking-up its products' retail prices. Were the firm to redistribute this profit (to its share or equity holders, or even to the general government sector in payment of taxes), the latter would ultimately be used up for the final consumption of period-1 yet unsold output. The fact that there exists an unsold output matching the newly formed profit
cannot be overlooked\textsuperscript{65}. If this profit is redistributed, no discrepancy between total supply and total demand could ever be detected in the economy, for such a gap would not really exist, as we hope to have established by now. It is therefore the firm's 'productive consumption' of (invested) profit that ought to be of interest to those aiming to study, and hopefully manage, actual inflationary pressures. So, let us focus on the new production set up by the firm's investment in period 2. By investing its profit on the factor market, the firm acquires a new capital good through the very same payment of its production costs (equal to 10 pounds in the example at hand). Assuming that the firm's production in period 2 gives rise to a total output measured by 100 money units (as resulting from the corresponding wage bill), 90 percent of the latter units are deposits of 'full' money. Still more precisely, 90 money units out of 100 have the power necessary, and sufficient, to purchase the whole stock of goods and services newly produced in period 2 (a stock worth 90 pounds because of the firm's purchase (10) on the factor market). There is nevertheless a remaining 10 percent of money units which, although created in the same income-generating finance process (to use Davidson's phrase), have no real object in terms of current saleable output. Investment of the firm's profit on the labour market removes in fact the corresponding output (10) from the newly formed deposits at the very instant workers obtain their remuneration, so that the measure of the new production remains numerically unchanged (100) but its actual object is reduced \textit{ipso facto} (to 90). In Schmitt's own words, '[t]his withdrawal leaves \textit{money} wages unaltered but decreases their purchasing power' (Schmitt 1984: 208, our translation)\textsuperscript{66}.

In sum, as we shall see more in detail in the next chapter, when a pre-existent income is consumed by an income-generating expenditure, the present payment

\textsuperscript{65} This output can be viewed as the wage-earners' forced saving. In fact, workers have no alternative but not to consume the part of total output corresponding to the income that firms capture on the product market by the mark-up of prices.

\textsuperscript{66} 'Ce prélèvement laisse les salaires \textit{monétaires} intacts; il ne diminue que le produit dans les salaires \textit{monétaires}' (Schmitt 1984: 208).

205
structure of any monetary production economy brings about an anomalous, and paradoxical, outcome. Although total supply and total demand are identically equivalent as measured by the wage bill, total demand implies a sum of empty money that elicits a pathological disequilibrium between money and current output. With respect to the previous example, this paradox can be restated in numerical terms as follows (figures relate to period 2).

- Total supply and total demand are both equal to 100, since 100 is the amount of the wage bill paid out to workers for the new production, in strict adherence to the principle of effective demand described earlier. The newly produced goods and services elicit a wage bill of 100, which is their precise measure and defines at the same time an identical amount of wage-earners' bank deposits.

- Total supply and total demand are not equal, because, at the very instant of its formation, output is curtailed by the firm's purchase occurring on the factor market. Receiving a sum of bank deposits numerically equal to 100, wage-earners really obtain overall only 90 units of purchasing power -- before any mechanism of income distribution may come into play on the product market.

Surprisingly enough, the disequilibrating exchange between the monetary flow and the real flow does not impinge on its actual outcome, that is, identical demand and supply in the economy as a whole. However, although total supply and total demand are equal in 'constant money' (that is, in terms of purchasing power) at any point in time, there is also an excess total demand in 'current money' (Schmitt 1972: 156; Gnos and Schmitt 1990: 68; Bradley et al. 1996: 133). Indeed, the numerical value of total demand (100) differs from its real value (90) because of the investment of the firm's profit on the labour market, an operation that ipso facto reduces the purchasing power of income holders at the instant the latter income is formed. The paradox of a disequilibrium within an identity is thus explained when it is shown that 'a part \( p(r) \) of output is sold before any expenditure of national income [on the market for
produced goods and services' (Schmitt 1988: 192); income and output being the two identical terms of the macroeconomic relation resulting from the new production process, where an accumulated profit equal to \( p(r) \) is invested.

Some authors have rooted this approach in Keynes's *Treatise on Money*, and in particular in his analysis of 'profit inflation' (Keynes 1930: Chapters 10-11)\(^6\). In the *Treatise*, in fact, Keynes seeks the explanation of inflation in the essence of capitalist economies, that is, in the investment of profit in the production process. More specifically, Keynes's message points out that 'profit inflation' originates in a new production of investment goods or, as he puts it, in a production of new investment goods (Keynes 1930: 135-138). As a matter of fact, by 'financing the production of investment goods with a pre-existing profit, the firms create an excess demand in real goods' (Bradley et al. 1996: 122). Restated in modern terminology, Keynes's path-breaking analysis indicates therefore that inflationary disequilibria do not arise simply because too much money chases too few goods, but fundamentally because a part of the newly produced output is definitively not for sale to income holders, although it is included in the measure of current national income.

Stated in its most troublesome terms, this conclusion amounts to the proposition that the result of the blending between income-generating and income-destroying expenditures on the factor market 'is that the goods corresponding to the net investment of profit are no longer in the economic possession of income holders' (Cencini 1996: 54). As Schmitt pointed out, it is because investment goods become the definitive property of 'a "non-person", [i.e.] the set of the country's disembodied firms' (Schmitt 1984: 208, our translation)\(^6\), that the emission of empty money leads to an irremediable monetary disequilibrium\(^6\). Were the new capital goods, that is, the

---


\(^6\) If the investment 'pays off', the output newly produced with the capital goods resulting from profit investment will be associated with new money units, as these units are issued in the new, periodic
goods produced via the investment of profit, entirely owned by 'embodied' firms (i.e. their share and equity holders), the economic system would not suffer the malignant disorder elicited by the present recording of capital accumulation in the banks' bookkeeping. In particular, any income that wage-earners are at present deprived of on the factor market would be distributed to other categories of economic agents, without such a distribution being a possible inflationary phenomenon (as recalled many times already). So much so that, at any point in time, shareholders and entitled agents would own the entire stock of accumulated capital$^{70}$, for which a one-to-one relationship between its financial and material aspects would absolutely exist in the economy.

It is particularly to the attempt at working out a structural monetary framework capable of achieving such a correlation between the two aspects of real capital that the next chapter will be devoted, after a more in-depth analysis of the consequences brought about by the emission of empty money just sketched here.

---

$^{70}$ Notice that ownership does not imply location. Share and equity holders can own instrumental capital which is located with firms, and serves their production plans.
CHAPTER 6
INFLATION AND CAPITAL ACCUMULATION

In Chapter 5 we observed that inflation is a macroeconomic disequilibrium, that is, it affects the relationship between money (total demand) and output (total supply). To acquire an understanding of the macroeconomic nature of inflationary pressures, we explored the production process, instead of focusing on the exchange of goods already produced. More precisely, we set off the investigation of profit investment on the factor market (i.e. capital accumulation) as an alternative direction for monetary research into the inflation problem. This chapter seeks to forge ahead this line of investigation, first put to the fore by Schmitt (1984). In particular, it attempts to point out that the most serious consequences of the emission of empty money on macroeconomic equilibrium occur when the capital accumulated in the whole economy is replaced within a monetary structure of payments not fully complying with the fundamental rules of non-inflationary bank money. This theoretical analysis is therefore not only fundamentally macroeconomic; it is also monetary, because it focuses on the financial capital invested in the means of production.

To introduce the reader to this macro-monetary analysis of inflation, let us say from the beginning that, in this framework, the problem of inflationary pressures is related to the process of accumulation in a capitalist economy, which in reality may occur in an orderly or in a disorderly way. To avoid any possible misunderstanding, let us stress that order and disorder refer to the manner in which accumulation of capital, that is, profit capitalisation, is recorded in the banks' bookkeeping. The agents' forms of behaviour are not at stake here. In other words, it is not the

---

1 Contrast the explanation based on the possible different forms of savings behaviour across the various sectors of the economy, or between different categories of income holders, as put to the fore by Harris (1978) in his analysis of Capital Accumulation and Income Distribution. A classification of growth and distribution theories based on various savings assumptions and on the role of technical progress is provided by Baranzini (1991: 45-48). See also Pasinetti (1993: 82-104) for a macro-
behaviour of the different categories of economic agents that can account, ultimately, for the disequilibrium between money and output existing in a national economy. As Cencini states, 'the relationship between money and output can be pathologically modified by a simple accounting mechanism that does not pay sufficient attention to the banking nature of money and to its functional link with production and circulation' (Cencini 1995: 70). In short, the accumulation of capital may take place in a bookkeeping structure of banks' accounts that mechanically respects the distinction between money, income and fixed capital (order), or in an accounting structure that does not (disorder). An example of structural monetary disorder has been explored in the preceding chapter, when we discussed the problem of an excess of bank credit (i.e. benign inflation). Centred on output circulation, that example showed that inflation may result from a too simple book-entry structure, which does not enable banks to distinguish between money emission and financial intermediation. Yet, a further, and more worrying, analytical case must be investigated, a case where the focus of analysis is on production rather than on exchange of commodities already produced. As Schmitt (1984) pointed out, (malign) inflation originates in the production process; in particular, in the production of amortisation goods destined to replace fixed capital. It is to this complex, and difficult, object of inquiry that we turn next, in an attempt to carry on the investigation of empty money started off in the last section of Chapter 5, and to forge ahead the Schmitt framework.

---

2 Theoretical analysis of consumption, savings and inter-temporal distribution of income in a pure labour economy.

2 In other words, the monetary recording of the accumulation of capital in the banks' bookkeeping may be inflationary or not. See the last subsection on this point.

3 See the second section of this chapter for analytical elaboration.
EMPTY MONEY AND INFLATION

As noted in the previous chapter, profit investment on the labour market elicits empty money. It is because the payment of the wage bill out of a pre-existent income entails a compensation of their amounts in banks' balance sheets that current period wage-earners are credited with a sum of empty money, that is, with nominal income deprived of purchasing power over current output. Reasserted in these terms, it should be clear that the emission of empty money cannot be imputed to any form of economic behaviour. In particular, the firms' investment of profit is not to blame, and the formation of the latter is indeed definitely legitimate. As has been stated by Cencini, ‘[t]he anomaly lies with a mechanism inappropriate to the nature of money, with a monetary structure still defective, and not with the decision to invest profit more or less productively’ (Cencini 1995: 89). Despite the very high banking standards existing in present-day payment systems, which aim to ensure both the efficiency and the stability of the whole financial system, there exists a yet unnoticed discrepancy between the actual working of the monetary structure of payments and the principles of modern money. More precisely, since the wage-earners' remuneration and the capitalisation of profits are recorded today in the same 'department' of the banks' bookkeeping, the pathological birth of empty money cannot be avoided in actual facts. As we shall see later, an improvement of the book-entry structure of bank accounting could eradicate the emission of empty money, leaving agents' forms of behaviour totally unaffected and completely free (subject to a set of moral values that any civilised society ought to have).

Yet, before putting forth some policy proposals aimed at improving the structural monetary framework of modern banking and payment systems, we need to refine the analysis of empty money. More precisely, we must further explore the link between empty money and inflation, which in the preceding chapter has just been proclaimed and not really treated yet.
Empty money, forced saving, and fixed capital

Let us start by recalling that profits are formed as consumers (i.e. households) transfer to firms on the commodity market an income which exceeds the production costs of the goods and services purchased. In fact, whilst the households' expenditure on the market for produced goods and services is income-destroying for the whole economy to the extent that it enables firms to cover factor costs of the output sold, it merely conserves income for the amount exceeding the latter costs. Indeed, as the abundant Post-Keynesian literature existing on this topic clearly shows, profits are elicited by the firms' mark-up of retail prices over factor costs. Now, it is a fact that any income captured on the commodity market entitles its holder to buy the corresponding goods (which define its objective substance), provided that the relevant production costs are entirely paid for. This is evident when profits are redistributed to, say, the firms' stockholders, who can thus consume what has not yet been purchased by original income holders (i.e. wage-earners). To focus on the point at issue, it must indeed be kept in mind that the firms' profit is defined by, and corresponds to, what may be labelled the wage-earners' forced saving. When retail prices are higher than the corresponding output's costs, the workers' income expenditure elicits a stock of wage goods which become in fact the firms' property. These goods literally are the substance of the wage-earners' forced saving of the current period. They define at one and the same time the real object of the monetary profit earned by firms on the product market. They thus represent the original form of capital, which can be invested in a new production process by 'productively consuming' the accumulated profit. As Schmitt puts it, '[c]apital goods, or investment goods, are not the primitive form of fixed capital; they are the secondary form, obtained by the "productive

---

4 See previous chapter (note 65).
consumption” of the wage goods saved in the formation of the monetary profit’ (Schmitt 1984: 170, our translation)\(^5\).

It is also possible to couch the same point in different terms, perhaps more familiar to the reader of the present work. By paying the wage bill out of pre-existent deposits (i.e. accumulated profits), firms become the owners of the output produced through the investment of profit. As already noted at the end of the previous chapter, this output is purchased on the factor market at the moment it is formed. In fact, the corresponding wage bill is deprived of its real object (i.e. the new investment goods) when it is paid out to wage-earners. The money units paid out to these wage-earners are literally and materially empty, because of the inclusion of an income-destroying expenditure into an income-generating expenditure (see Figure 5.1).

Yet, this complex operation is not inflationary in and by itself. In other words, the emission of empty money originating in the net investment of profit (on the labour market) does not give rise to a disequilibrium between total demand and total supply in the economy as a whole. In fact, as we have noticed, the formation of the firms’ profit elicits an identically equivalent stock of goods, which are the wage-earners’ forced saving. So much so that the monetary profit gained by firms on the product market has the power to purchase the goods yet unsold to wage-earners (or, more generally, to initial income holders). Hence, when the investment of profit occurs, the resulting sum of empty money has the purchasing power that corresponds to the wage goods accumulated when profit has been formed. Profit investment is, ultimately, the conversion of a stock of wage goods into a ‘stock’ of new capital goods (Schmitt 1984: 166)\(^6\). Deprived of the power to purchase the newly produced investment goods


\(^6\) We put the word ‘stock’ in inverted commas, because the newly produced capital goods are not for sale to income holders. The firm owns them since their very formation, because the latter results from the investment of profit (i.e. its capitalisation). This does not exclude, however, both inter- and intra-industrial transactions (see Bradley and Guos (1991: 185-186) and Friboulet (1991)).
(because these goods are definitively acquired by firms on the factor market), the current period wage-earners thus have nonetheless the power necessary, and sufficient, to buy the accumulated stock of consumption goods which are the original form of retained profit. In Schmitt's own words, 'the money units issued in the investment of profit find ultimately a “body”, that is, the wage goods saved within the process by which monetary profit is formed' (p. 190, our translation)^.

Let us try to reconsider this process through a numerical example. In this analytical context, the period (1) when profit is formed may be distinguished from the period (2) when profit is invested and generates empty money. Suppose that the firms' mark-up pricing policy in period 1 allows them to earn a profit equal to 10 percent of total households' current expenditure on the goods market. Accordingly, if original income holders spend an income of 100, the firms' profit is equal to 10. This means that households consume an output of 90 (because 90 percent of their expenditure covers the production costs of sold output), and transfer an income of 10 to firms, which then have the power to purchase the equivalent output. Now, since profit is not consumed but invested, the latter output is not sold to period-1 income holders. It defines indeed period-1 forced saving, as we have already recalled. As a matter of fact, this profit is spent on the factor market, where it purchases labour and, consequently, its result. We need therefore to focus on a new production process, namely on the production of new capital or investment goods taking place, say, in period 2. When the latter production occurs, as we know, workers obtain a sum of empty money to the extent that the wage bill is paid out of pre-existent deposits.

^  "[L]a monnaie émise dans l’investissement du profit trouve finalement un “corps”, en l’espèce des biens-salaires épargnés dans l’acte de la formation du profit monétaire” (Schmitt 1984: 190). Note that the sale of this stock of wage goods may, of course, give rise to a new profit for the seller. The point to underline here is that the emission of empty money in the investment of profit is not yet harmful for the economy, since the corresponding money units acquire a purchasing power over the stock of wage goods saved when profit has been formed. The actual distribution of this purchasing power (between firms and households) is not germane to the problem at stake, and will therefore be put aside in the present analysis."
Remunerated with the firms’ accumulated profit (equal to 10 in our stylised example), wage-earners literally obtain a non-income, or a sum of money units with no purchasing power at all over period-2 output. However, the stock of wage goods corresponding to the period-1 forced saving (necessarily equal to 10 in the case in point) is still available, for sale, in period 2. Hence, the amount of empty money generated by the firms’ investment of profit acquires, ultimately, a real object. It has precisely the power to purchase the goods saved in the period when this profit has been formed. In a nutshell: ‘Paid by profits, workers obtain an empty income and, as a compensation, a claim over the wage goods saved by the previous transformation of wages in profits’ (Schmitt 1984: 215, our translation).

On the whole, neither the formation of profit nor its ‘productive consumption’, that is, net investment, can affect monetary equilibrium. When profit is formed, total demand is identical to total supply, for no redistribution of income can impinge on the relationship between money and output (see Chapter 5). On the other hand, as we have seen earlier on, when profit is invested, the amount of empty money elicited by the remuneration of workers through an income-destroying expenditure on the factor market ultimately does not generate an inflationary gap in the economy either. To the extent that the wage goods accumulated when profit was formed take the place of the newly produced capital goods definitively acquired by firms, the current period wage-earners obtain in fact a money capital, that is, the primitive form of capitalised income. The monetary disorder called forth by empty money is thus not much worrying yet, because current period wage-earners automatically own a capital -- in the form of the wage goods previously produced -- in substitution of the income they were entitled to hold as a result of their remunerated activity in the investment goods sector. ‘It remains true that period-2 production elicits empty money, but the latter is

---

inoffensive because the capital goods instantaneously withdrawn from real wages [...] are exactly compensated by the wage goods [previously] saved' (Schmitt 1984: 205-206, our translation).

Now, the presence of capital goods (e.g. machinery) in the economy requires their maintenance. Introduced by Keynes's "Appendix on User Cost" (Keynes 1936: 66-73) -- which the author of The General Theory did not have the time to develop further --, the macroeconomic investigation of fixed capital amortisation may be fundamental for the analysis of inflation. In so far as attention is focused on net investment, no disequilibrium between money and output can indeed be discovered in the whole economy, at any point in time (for, let us repeat it, the newly created empty money instantaneously acquires as its substance the wage goods saved when profit was formed). By contrast, as Cencini and Baranzini state, ‘the production of amortisation goods plays a determinant role in explaining the possible discrepancy between total demand and total supply’ (Cencini and Baranzini 1996: 8). It is to the analysis of this production that we must now turn, in an attempt to explore the inflationary gap further and deeper. To be sure, the ultimate origin of (malign) inflation has yet to come to light. Our study has then to be carried on, if it aims to understand the efficient cause of the loss of money's purchasing power.

The amortisation of fixed capital

As is well known, Keynes (1936) defined by ‘user cost’, $U$, the cost incurred for the use of fixed capital and the production of replacement goods, that is, amortisation. To be true, as pointed out by Chick (1983: 49-50, 91-92), Keynes’s definition of user cost was broader than that, since it also included wear and tear of fixed capital not yet

---

9 'Il reste vrai que la production de $p^+$ comprend une émission vide, mais elle est inoffensive puisque les biens-capitaux instantanément retirés des salaires réels [...] sont exactement compensés par les biens-salaires épargnés' (Schmitt 1984: 205-206, translation fitted to match our numerical example).
replaced and 'optimal maintenance', a forward-looking approach which we do not consider here. For the sake of argument we shall include in $U$ only the production costs of replacement goods. In fact, this seems also to be the case for the bulk of existing literature.

Now, a quick look at the scarce literature on user cost shows that contemporary authors are not at ease with the macroeconomic dimension, and implications, of this difficult, and multifaceted, concept (see e.g. Torr (1992) and the references therein). 'While it is rare to find a comprehensive discussion of user cost in the post-1936 macroeconomic literature, it seems fair to say that those expositions that have appeared have tended to concentrate on clarification of what Keynes meant by user cost rather than on development of the concept' (Torr 1997: 144). In particular, within the economics profession there seems to be unanimity for considering total output of an economic system net of aggregate user cost (see Keynes (1936: 67)). To take only a recent, but important and authoritative, example, let us quote three significant passages from A 'Second Edition' of The General Theory, whose Chapter 8 is precisely on "User Cost". Writing 'as J. M. Keynes', the author of the chapter states a generally undisputed point, namely that replacement goods ($R$) are nothing more than intermediate goods$^{11}$, because they replace the fraction of fixed capital used up over the period under examination. From a purely rhetorical standpoint, the three quotes form an increasing climax:

User cost has one footing resting in microeconomics and the other in macroeconomics. [...] At the macro level, user cost must [...] be subtracted from total sales, since it is the cost of items used up in the production process and as such must not be included in the total of final consumption and investment goods (Torr 1997: 133).

User cost must always be incorporated in the analysis, at both the micro and macro levels. [...] We take it into account at the macro level to ensure that our

\[ 10 \text{ For an illustration of recent attempts at clarification, see Davidson (1987) and Deprez (1993).} \]

\[ 11 \text{ Conversely, consumption goods (C) and investment goods (I) are final goods.} \]
measure of total output does not include goods used up in the production process (p. 134).

At the macro level we need to subtract user cost to arrive at a measure of the aggregate level of output and income. [...] If we do not do so, national income will include those goods used up in the production process.

To arrive at the income for the whole economy, we need to take the total sales of all firms and subtract certain intermediate sales. User cost identifies which inter- and intra-firm transactions must be subtracted (p. 139).

In practice, if not in theory, most economists assert that the measure of output as a whole, hence of national income, can be arrived at 'either by the method of subtraction \(Y = A - U\), where \(A\) is the sum of total sales of all firms, and \(U\) is total user cost) or by the method of addition \(Y = C + I\)' (Torr 1997: 140, brackets' content rendered explicit for the clarity of exposition). This statement is further reinforced by the allegation that '[t]he method of addition amounts to the same thing as the method of subtraction' (p. 134).

Now, 'the method of subtraction' does not pertain to macroeconomic thinking. In this framework, as we have noted in the previous chapter, no link can be conceived of between \(Y\) (production) and \(A\) (exchange): national income is smaller than \(A\), for it includes only production costs (Schmitt 1972: 109). In fact, the sum of total sales of all firms includes, by definition, the whole amount of profits, i.e. that part of national income that firms can capture on the goods market by selling their products at a mark-up over factor costs. Further, since user cost has not to be mixed up with what Keynes (1936: 53) defined as 'prime cost', \(U\) cannot be thought of as the analytical (or even merely arithmetical) difference between \(A\) and \(Y\). We are thus left with 'the method of addition', if we really aim to avert any double counting. Yet, if all double counting must be disposed of, attention has also to be paid to counting everything. So, is it fundamentally correct to claim that the measure of national income, and hence of output as a whole, can be apprehended by the canonical expression \(Y = C + I\)? Still less enigmatically, from a macroeconomic vantage point would it not be necessary to write that \(Y = C + I + U\), where \(U\) is the total value of replacement goods? As a
matter of fact, to exist, amortisation goods have to be produced by current-period workers\textsuperscript{12}. Ultimately, then, remuneration of the latter adds to the wage bill paid out by firms, thus counting for the determination of national income. As Schmitt points out,

\begin{quote}
[i]n fact, the production of the goods which replace the fraction of fixed capital that is lost in the period yields goods that are final and not merely, as is generally thought, intermediate. [...] ‘[R]eplacement goods’ exist in two distinct forms or definitions; this amounts to saying, curiously enough, that these goods belong simultaneously in the separate categories of intermediate and final goods (Schmitt 1996b: 77).
\end{quote}

As any economist or accountant knows, amortisation goods are intermediate goods, since they replace the fraction of fixed capital used up in the production process. This is also the empirical stance of entrepreneurs and, as such, it can be studied according to a microeconomic theory of the firm (see Davidson’s (1987) dictionary article for references along that line). However, replacement goods are also final goods, since their production costs elicit a wage bill which logically enters into the macroeconomic measure of the output newly produced\textsuperscript{13}. ‘Replacement capital must be newly produced in each period; there is no difference in this respect between R-commodities [where R stands for replacement] and the other final goods, produced for consumption and investment’ (Schmitt 1996b: 95).

If this analysis is correct, then the value of replacement goods \((U)\) has to be added to the value of \(C\) and \(I\) in national income determination, opening up a new avenue for

\begin{footnotesize}
\textsuperscript{12} As Schmitt observes, ‘[a situation] where the national income is equal to \(C + I\), could only exist if a national economy produced goods without the help of any fixed capital, or, at least, if no user cost were incurred in a given period by an economy. Clearly, such an economy cannot be found in the world as anyone can observe it, unless the chosen period is unduly, not to say ludicrously, short’ (Schmitt 1996b: 79).

\textsuperscript{13} A phraseology that might be more convincing for at least some readers relies on a Marxian interpretation of fixed capital amortisation, a process Wray attempts to render along Keynesian lines in the following terms: ‘constant capital used in production that is not replaced does not set any live labor in motion (to produce replacement means of production); in other words, it does not generate a wage bill (in the period it depreciates) in the means of production department (“investment” sector). Only replaced dead labor can lead to the realization of value by creating wages and spending on consumption goods’ (Wray 1999: 14).
\end{footnotesize}
both macroeconomic theory and research. More specifically, to focus on the problem of this thesis, the analysis of inflation takes on an entirely new dimension, since it has also to investigate the precise effect of fixed capital amortisation on the relationship between money and output. In this analytical framework, it may ultimately be possible to highlight the fundamental origin of inflation, which, as Schmitt maintains, "is not a purely monetary disorder; it is a pathology acting on money in its relationship with output: it thus affects both money and output' (Schmitt 1984: 506, our translation). In particular, the worrying pathology of (malign) inflation can be discovered by a macro-monetary analysis of the production of replacement goods, i.e. the monetisation by banks of fixed capital amortisation. Certainly this is a topic which needs to be further investigated by our profession. Indeed, to the best of our knowledge, the monetary issues of fixed capital amortisation are still neglected in contemporary macroeconomic analysis. There thus seems to be a strong case for further research in this field.

Without entering here into the technical details of the Schmitt analysis, it must be emphasised that, according to the line of thought we adhere to, the inflation problem originates in a structural monetary disorder that up to now has framed the recording of fixed capital accumulation in the banks' bookkeeping. The macro-theoretical analysis developed by Schmitt (1984) and Cencini (1995, 1996) shows indeed that the

---

14 'L'inflation n'est pas un désordre purement monétaire; elle est une pathologie définie sur la monnaie dans sa relation avec le produit: elle touche donc à la fois la monnaie et le produit' (Schmitt 1984: 506).

15 As clearly noticed by Joan Robinson, the problem of fixed capital replacement is difficult to analyse, at least in macroeconomic terms, also because it is not straightforward to distinguish net investment from amortisation of existing capital goods. 'New investment is normally going on and at the same time the composition of output and techniques of production are changing. A worn-out plant is rarely replaced by an exact replica of its original self, and when the physical specification of the replacement, or its expected future life, is different, or the market conditions in which it will operate are different, there is no precise and unambiguous criterion by which to judge whether it is exactly equivalent to what it replaces. We cannot then draw a clear line between replacements and new investment, though firms must adopt accounting conventions to make the distinction between amortisation and profit' (Robinson 1956: 42).

production of amortisation goods gives rise to a set of book entries recorded in a still imperfect structure of bank accounting, that is, a monetary structure not yet fully in line with the fundamental requirements of bank money. It is because the amortisation of fixed capital is recorded in a yet too simple accounting structure of the banking system that a disequilibrium between total demand and total supply can occur in the economy as a whole. Eliciting a sum of empty money, in present-day capitalist economies the production of replacement goods through the investment of profit increases total demand without identically increasing total supply. Whereas in the net investment of the firms' profit a pre-existent stock of commodities can ultimately ‘fill the gap’ -- thus preserving the overall monetary equilibrium in the national economy --, fixed capital amortisation is at present irremediably pathological, for total demand is definitively greater than total supply\(^{17}\). ‘It is appropriate to say \textit{irremediably} since, by contrast to the empty money elicited by net investment, the empty money obtained from the production of amortisation goods results in a sum of money units whose emptiness is not compensated at all; this time, no pre-existent wage goods are waiting to fill up the gap’ (Schmitt 1984: 223, our translation)\(^{18}\).

In sum, the pathological nature of fixed capital amortisation is the concrete, and mechanical, result of a still partial discrepancy between the actual structure of the banks' bookkeeping and the essence of bank money, a discrepancy that elicits a sum of nominal income deprived of value, thus epitomising an excessive money supply. All in all, inflation does not occur because, as the (in)famous phrase goes, too much money chases too few goods, but because the relationship between money and output

\(^{17}\) Note that the proposed solution is not to allow fixed capital to run down (to zero). It consists in a structural monetary reform, as we shall see later.

\(^{18}\) \textit{‘Il convient bien de dire irrémédiablement car, à la différence de l’émission vide définie par l’investissement net, l’émission vide induite de la production des biens d’amortissement aboutit à une monnaie dont la vacuité n’est nullement compensée; cette fois, aucune épargne de biens-salaires n’est en attente pour remplir le vide’} (Schmitt 1984: 223). As noted at the end of the previous chapter, since fixed capital is appropriated by the set of the country’s ‘disembodied’ firms, its amortisation (i.e. the production of replacement goods) does not elicit a stock of goods to be sold to ‘embodied’ agents.
established by the banks' monetisation of current production is instantaneously hampered by capital amortisation. To be more precise, the actual monetary structure of payments is such that the production of replacement goods gives rise to an output instantaneously withdrawn from the set of saleable commodities, although the corresponding nominal wage bill adds to the sum total of newly formed income. To put it in the terms of Chapter 5, although total demand is identically equivalent to total supply in constant money, the former is greater than the latter in current money. There is an irreducible excess of demand on the product market because of fixed capital amortisation.

AN OUTLINE OF STRUCTURAL MONETARY REFORM

In this section, we shall attempt to polish up the solution of inflationary pressures put forth by Schmitt (1984, 1995-1996) and Cencini (1996), who insist in particular on the necessity to spread monetary operations over three bank departments, functionally distinct in bookkeeping terms, to have an inflation-proof economy. In short, a structural reform of the bookkeeping framework within which banks operate is the *sine qua non* condition to make sure that any monetary transaction actually complies with the nature of modern money -- independently of the agents' behaviour, let us stress it again. In other words, to avoid the generation of empty money, the working of the banking system must conform 'empirically' to the fundamental distinction existing between money, income and capital, so neatly articulated by Cencini (1995: 70-74; 1999: 139-142). For the sake of exposition it is worth treating the case of benign inflation first, so as to follow an order of increasing difficulty and to provide an explicit link with the subject matter of the last section of Chapter 5. The solution of malign inflation relies indeed on the same principles, and is ultimately a refinement, of the reform treated next.
On the book-entry distinction between money and credit

Being the direct result of the banks' monetisation of the production process, excessive money supply (that is, empty money) might be thought of as depending on economic behaviour of the banking system as a whole. In Chapter 5 we observed indeed that bank credit may be granted beyond the level established by the relationship between money and current output. We also attempted to show, however, that credit-led inflation can never be so troublesome for the macroeconomy, since it entails a self-correcting process, namely credit reimbursement. Yet, to set the record straight, we may begin by addressing the issue of making sure that the emission of money is not confused, most importantly in practice, with a credit operation originating in banks. Still more precisely, since excessive bank loans originate in the 'empirical' mixing up of money and credit, the solution consists in introducing a clear-cut, and operational, distinction between money emission and financial intermediation in the banks' bookkeeping.

As Ricardo clearly noted at the outset of his 1823 Plan for the Establishment of a National Bank, published in 1824 six months after his death, banks still perform two fundamentally distinct operations in the economy: they issue money and make loans to the non-bank public. Let us quote Ricardo:

The Bank of England performs two operations of banking, which are quite distinct, and have no necessary connection with each other: it issues a paper currency as a substitute for a metallic one; and it advances money in the way of loan, to merchants and others.

That these two operations of banking have no necessary connection, will appear obvious from this, -- that they might be carried on by two separate bodies, without the slightest loss of advantage, either to the country, or to the merchants who receive accommodation from such loans (Ricardo 1824/1951: 276).

Although applying his analysis to the central bank only, Ricardo correctly pointed out that the emission of money has neither conceptually nor in practice to be mixed up with the granting of credit by banks. As Keynes was also to distinguish in early drafts
of his *Treatise on Money* (Moore 1988: 195), the money-purveying function must be kept distinguished from the credit-purveying function, both in theory and practice. However, Ricardo certainly went too far when he claimed that these two functions bear ‘no necessary connection with each other’. In fact, the creation of modern money requires that banks grant a credit to the economy. This complex operation, that is, a payment, involves money’s creation and bank lending in such a close relationship that the infelicitous expression ‘credit-money’ has often been used to stress that link. ‘Yet, from an analytical point of view, these two operations are completely distinct: by creating money banks do not lend positive income to the public but a simple promise which becomes a net credit only when it is spent by its borrower’ (Cencini 1988: 65). In other words, as so often claimed by adherents to the so-called ‘monetary circuit approach’, the emission of money is the banks’ spontaneous acknowledgement of debt to the economy. In this framework, Parguez and Seccareccia have recently noticed that ‘money emerges always as a “debt” (or liability) issued by this third agent [i.e. the bank] on itself’ (Parguez and Seccareccia 2000: 101). By this very action banks indeed do not lend anything, either real or nominal. They actually create money, that is, the numerical counter of any monetary transaction (as we have seen in Chapter 4). ‘No creation is an intermediation. The reciprocal is also true: no intermediation is a creation. In fact, if banks lend an amount they borrow, they create nothing’ (Schmitt 1984: 303, our translation). So, the book-entry distinction between money and credit is designed to put into banking practice the conceptual, and fundamental, distinction between money’s emission and financial intermediation.

---

19 As pointed out by Schmitt, ‘[m]oney and payments are one and the same thing. No money, if correctly defined, exists either before or after a given payment’ (Schmitt 1996b: 88).

20 ‘Aucune création n’est une intermédiation. La réciproque est vraie aussi: aucune intermédiation n’est une création. En effet, si les banques prêtent une somme d’argent par elles empruntée, elles ne créent rien’ (Schmitt 1984: 303).

21 Vallageas (1988: 189) claims that this distinction is impossible on account of the variety of bank loans.
Keeping within the limits of an outline, let us try to show how the creation of money and the granting of credit could be distinguished operationally in the banks' double-entry system of accounts. This distinction is not merely academic, as some practitioners might think at first, but has a direct, and important, implication on what, following Schmitt's (1984) work, we dubbed 'benign inflation', that is, the possibility for commercial banks to grant credit in excess of the existing deposits of the public.

To begin with, Table 6.1 is an attempt to illustrate the structure of the banks' bookkeeping that can avert credit-led inflation, although we noted in the previous chapter that this kind of inflation does no irremediable harm, and may even be pretty useful to the economy in terms of growth and output sale\textsuperscript{22}. Notice that the example at hand focuses on the payment of current production costs, to underline that any kind of loan implies an income -- which is the result of economic activity and not of a mere stroke of a bank's pen\textsuperscript{23}.

\textbf{Table 6.1} \hspace{1cm} \textbf{The bi-partition of the banks' bookkeeping\textsuperscript{24}}

<table>
<thead>
<tr>
<th>BANK</th>
<th>Monetary department (I)</th>
<th>Financial department (II)</th>
</tr>
</thead>
<tbody>
<tr>
<td>liabilities</td>
<td>assets</td>
<td>liabilities</td>
</tr>
<tr>
<td>(1) Department II</td>
<td>100</td>
<td>Firm</td>
</tr>
</tbody>
</table>

\textsuperscript{22} The choice to avert credit-led inflation might therefore be considered as an issue of economic policy. The decision of economic policy makers notwithstanding, from an analytical point of view this problem has however to be investigated on purely conceptual grounds.

\textsuperscript{23} In Table 6.1 we assume for the sake of simplicity that a single bank represents the banking system as a whole. It need not be emphasised that the proposed bookkeeping reform applies to any bank existing in the real world.

\textsuperscript{24} Notice here the similarity with the 1844 Bank Charter Act, which followed Ricardo's \textit{Plan} some twenty years after it was put forth and separated the business of the Bank of England into a 'Department of issue' and a 'Department of deposit and discount'. See Rotelli (1982: 242-250) for more details on the 1844 Act.
The first department, I, is in charge of money emission. The second department, II, records the newly formed income (which is in the form of a bank deposit). Referring to the payment of the wage bill, department I must intervene to issue the monetary form of the transaction (say, 100 money units) -- as asked by the payer (a firm in the case in point) --, once the credit-worthiness of the latter is considered acceptable by the bank’s managers (entry (1)). Department II records instead the purchasing power generated by the monetisation of current production, which defines at one and the same time a new money income whose initial holders are workers, as recalled several times already (entry (1')). The accounting rule is universal, and its implementation mechanical:

- For each payment, the payer is entered into the monetary department, for the bank issues a spontaneous acknowledgement of debt to ‘count’ the object of the monetary transaction.
- Simultaneously, the payee is entered into the financial department, since he or she holds a bank deposit, that is, an income saved in the form of a drawing right over current output.

Hence, monetary and financial operations always go together. They are in fact the two faces of the same reality, as we attempted to show in the two previous chapters and, in much greater detail, in Rossi (1998). The structural connection between the first and the second department in the banks’ bookkeeping is therefore entirely in line with the nature of modern money. Without repeating the whole analysis elaborated in Chapters 4 and 5, let us merely recall here that wage-earners own an income (hence

---

25 Ricardo himself used to think in terms of departments in his Plan (Ricardo 1824/1951: 291).

26 As put by Schmitt, ‘[a] computer program can be devised and established whose theoretical goal and practical effect is, when applied by the banking system of the nation, to eradicate all positive creations of purely nominal incomes’ (Schmitt 1996b: 105). Applied to the practical distinction between money emission and financial intermediation, the required structural monetary reform is therefore a matter of mechanically preventing the mixing-up of two kinds of fundamentally distinct operations.
the qualifier income holders, IH) whose purchasing power is defined, exactly and objectively, by the newly produced output27 (we abstract here from the pathological emission of empty money, which will be taken up again in the next subsection). Thus then, workers do not own a mere sum of money proper when they are remunerated for their labour services. The structural distinction between department I and II is in fact the translation in modern banking of the fundamental distinction between vehicular money and bank deposits (see, for instance, Figure 4.4). It also is the only operational way to distinguish a purely numerical form, issued by banks, from its real content, which must necessarily be provided by human effort (i.e. production)28.

So, let us now consider the time-dimensional sequence of bookkeeping entries epitomising the financial operations involved in the case of (say) a second-hand transaction in housing assets. This will enable us to provide an explicit link with the stylised example introduced while investigating credit-led inflation, in an attempt to show the fundamental difference when money emission and financial intermediation are kept separate in the bank's accounts (Table 6.2).

27 Remember that this does not mean that wage-earners can obtain the whole output newly produced when they spend their income. In fact, firms may mark-up retail prices, so as to capture on the product market a share of national income (formed on the factor market, when workers are paid).

28 Present bank accounting does not distinguish the numerical means of payment from the very object of the latter, that is, income in the form of bank deposits. See Rossi (1998: 35-43) for an attempt to introduce this distinction in the actual structure of the banks' bookkeeping.
Entries (1) and (1’) have already been explained above; they are simultaneous and epitomise the dual nature -- monetary and financial -- of any payment. Entries (2) and (2’) define the transformation of the firm’s monetary debt into a financial one by, say, the bank’s end-of-day settlement protocol; they are recorded in the same instant, not necessarily contemporaneous with the first line of double entries29. ‘Cancellation of first department entries derives from the purely vehicular nature of money, whose “mark” is financial and can be found in the bank’s second department’ (Cencini 1995: 75). Entry (3’) is the result of entries (1’) and (2’): it shows that workers own an income (bank deposit) as a result of their remunerated activity by the firm, which has an identical debt for the financing of its newly produced stocks30.

---

29 One might admit that the second line of book entries has to be recorded by the end of a bank’s business day, when interest charges are computed in practice. However, from a theoretical viewpoint the first and the second double-entry lines can be made to coincide in chronological time, so that entry (3’) would be the immediate result of any payment (where workers (income holders) and the firm stand, respectively, for the payee and the payer). Today, only entry (3’) is recorded in bank accounting, since the latter does not (yet) distinguish between monetary and financial operations.

30 Since we focus here on the macroeconomic result of current production, we abstract from any pre-existent deposit the firm might have. See the next subsection for analytical elaboration.
Now, what happens if an agent whatsoever needs a bank loan to buy a second-hand item, say an housing asset? At first thought, the bank is able to accommodate any loan request (upon provision, needless to say, of adequate credit-worthiness of the borrower), since up to now the only requirement for a bank to grant a loan has been to respect the 'golden rule' of double-entry bookkeeping: any entry on one side of the balance sheet has to be balanced by an equivalent, and simultaneous, entry on the other side (see Chapter 5). It is here that the reformed structure of the banks' bookkeeping can avoid excessive credit to be granted, by separating the money-purveying function from the credit-purveying function. As pinpointed by Schmitt, 'without this separation, banks do not know the amount of savings deposited with them' (Schmitt 1984: 311, our translation). Indeed, as noted in the previous chapter, credit-led inflation is the result of an excess of bank loans with respect to the amount of savings. To repeat, any single bank can lend today an amount it literally creates, because of the lack of distinction between money and income (and, therefore, between money emission and financial intermediation) in our banking systems. The above structural bi-partition of the banks' bookkeeping is therefore necessary to make sure that the working of present-day monetary systems is consistent with the nature of modern money. 'The division of banks into two departments is thus very important to avert inflation. After the reform, each bank will know every day the exact amount it

31 Note that if this unspecified agent were to ask for a consumption loan to purchase the goods newly produced and stocked within the firm, the bookkeeping entries would be simpler than in the case at hand. To wit, entry (4') would be replaced by entry (6'), and entries (6), (7) and (7') would not exist, leaving the overall result unchanged, as in entry (8').

32 In A Treatise on Money, Keynes noticed that there is no limit to the amount of bank lending if banks 'move forward in step' (Keynes 1930: 26). Indeed, this unlimited growth of bank loans is the theoretical result of the accounting equivalence of assets and liabilities of any individual bank. Any credit granted ex nihilo by a bank is in fact always matched by an equivalent, and simultaneous, deposit, made by the recipient of the payment the bank makes on behalf of its client (the borrower). At present, it is only by the clearing mechanism that each bank can know its financial situation in the interbank market, that is, if its loans have been greater, equal to, or less than the savings deposited in it over the business day just closed. This is ex post information, as opposed to ex ante knowledge of the exact amount of savings.

33 'À défaut de cette séparation, les banques ne connaissent pas le montant des épargnes constituées chez elles' (Schmitt 1984: 311).
can lend, as is the case for a Savings and Loans fund. It will never be possible that the banking system as a whole can lend more than the savings it actually has’ (Schmitt 1995-1996, vol. II: 33, our translation)\textsuperscript{34}.

In fact, only the income temporarily saved within the whole economy (which in the above example amounts to 100) may be lent to borrowers, if any anomalous alteration in the relationship between money and current output is to be prevented. So, to refer to the case at hand, if (and when) this loan is granted by the bank, the payment of the second-hand item gives rise to entries (4) and (4')\textsuperscript{35}: entry (4) epitomises the fact that the payer has to be provided with the monetary form in which the payment is made; entry (4') records the purchasing power earned by the seller (of the house), who obtains a claim on a bank deposit in exchange for the housing asset he or she previously held. Entries (5) and (5') are then the transformation in the bank’s bookkeeping of the borrower’s monetary debt into a financial one. They depict the fact that, to obtain a loan from the bank, the borrower has to relinquish a financial claim (see previous chapter), which is recorded on the assets side of the bank’s financial department.

Now, supposing that the seller of the housing asset decides to spend his or her deposit -- recorded in entry (4') -- on the product market, entries (6)-(6') and (7)-(7') are recorded as soon as the consumption of current output (worth 100) takes place. Entry (6) defines the emission of the ‘numerical counter’\textsuperscript{36} of this operation, whose mark is financial and is recorded as entry (6'):: the firm obtains a claim on a bank deposit (100) for the sale of its total output (at a price of 100) to the seller of the house. Entries (7) and (7') transform the monetary debt of the payer into a financial debt of the seller.

\textsuperscript{34} ‘La division des banques en deux départements a donc une grande importance au titre des remèdes de l’inflation. Après la réforme, chaque banque connaîtra chaque jour les sommes exactes qu’elle peut prêter, comme c’est le cas pour une Caisse d’épargne. Il ne pourra plus arriver que la somme des banques prête au-delà de ses ressources’ (Schmitt 1995-1996, vol. II: 33).

\textsuperscript{35} Recall that unused overdraft and credit lines do not appear in the bank’s balance sheet.

\textsuperscript{36} The redundancy (numerical counter) is deliberate.
one, since, as explained above, any purchase of goods, services, or financial assets involves also a purchasing power and not merely a sum of numerical units. A purchasing power (in the form of a bank deposit) of 100 is therefore literally destroyed in the purchase of the very output defining its physical object\textsuperscript{37}. Indeed, we should not lose sight of the fact that all the income originally held by wage-earners is automatically lent to some other agents (a single borrower in the case in point), who ultimately spend it in place of its initial holders. As Keynes noted in his speech before the House of Lords on 18 May 1943, ‘[i]f an individual hoards his income, not in the shape of gold coins in his pockets or in his safe, but by keeping a bank deposit, this bank deposit is not withdrawn from circulation but provides his banker with the means of making loans to those who need them’ (Keynes 1980: 273)\textsuperscript{38}. It would in fact be naive to deny that ‘[b]anks are intermediaries in financial transactions that the depositors are not necessarily aware of’ (Gnos 1998: 46). Applied to the stylised example we are investigating, this amounts to saying that the original depositors (i.e. workers), willingly or not, exchange their drawing right over current production (entry (3')) with an interest-bearing financial claim over an identically equivalent future production (entry (8')) -- an issue we already drew attention to in the previous

\textsuperscript{37} This action closes the current period; a new production is necessary to open next period and to form new income. Note again that the retail price paid on the product market by, say, the seller of the housing asset may include the firm’s mark-up over factor costs. In this case, an income of 100 is only partially destroyed, to wit, for an amount equal to the firm’s production costs of the output sold. Indeed, entries (6') and (7') testify that an income expenditure of 100 enables the firm to reimburse its initial debt (recorded as in entries (1) and (2')) either to the bank or to its own wages fund.

\textsuperscript{38} No income holder can spend his or her income at the very instant when the latter is formed. Each income holder thus saves his or her income up to the instant (of their choice) when he or she spends it on the product market. The act of saving of the newly formed income by its original holder implies, therefore, that this amount is lent by the bank where it is recorded as a deposit. ‘The association between money and output elicits an income [in the form of] a bank deposit. And it is precisely because it is formed as a bank deposit that income is immediately lent. Through the financial intermediation of banks, savings are instantaneously lent by their initial owners and spent by their borrowers’ (Cencini 1995: 71). In The General Theory, Keynes clearly observed in fact that monetary transactions are of a bilateral character. He explained that ‘regarding an individual depositor’s relation to his bank as being a one-sided transaction, instead of seeing it as the two-sided transaction which it actually is’, is ‘an optical illusion’ (Keynes 1936: 81).
What the overall book-entry result (8') shows, in fact, is that original income holders have in the end a purchasing power defined by the financial assets relinquished by the borrower. In simple terms, this means that workers (i.e. IH) have a financial capital as the form in which their savings exist until they decide to spend them. 'Income is thus transformed into capital, and it is as such that money can play the role of bridge between present and future so clearly enunciated by Keynes' (Cencini 1995: 71).

To cut short the story, the structural bi-partition between the monetary and the financial department is necessary to avert an excess of demand on the market for current output. 'Before the reform of the banks' bookkeeping, that is, in present-day economies, the financial market can let (the sellers of bonds) obtain an income whose expenditure relates to another period's product; in this case, demand is excessive since pre-existent incomes are spent together with new income' (Schmitt 1984: 312, our translation). Let us try to illustrate this crucial point by referring to the numerical example made earlier on. In the present situation, i.e. when banks do not (and cannot) distinguish money from income, an excess of credit leads to an excess of demand for current output, because (say) 110 units of money chases an output of 100 (as measured by the total wage bill its production did generate in the current period) (see Table 5.2). In other words, 10 units of money may be lent ex nihilo to the non-bank public, in what is ultimately an advance of a future income via the banking system (see previous chapter). Here lies the problem indeed. Being a future income, its expenditure on the product market must be made to coincide with the final purchase (that is, consumption) of the very output it will define when the corresponding wage bill is paid out. This is not what happens today: as is the case in Table 5.2, a future income

---

39 'Avant réforme de la comptabilité des banques, donc dans l'économie actuelle, le marché financier peut faire parvenir (aux vendeurs de titres) des revenus dont la dépense est rapportée au produit d'une autre période; dans ce cas, la demande est excédentaire parce que des revenus anciens sont dépensés en concours avec les revenus nouveaux' (Schmitt 1984: 312).
income of (say) 10 -- advanced via the banking system -- is spent jointly with a current income of 100 for the acquisition of a current output worth 100, thus eliciting an excess of demand whose effect on prices is well known in modern capitalist economies. To wit, retail prices are made to rise to re-establish, in numerical terms, the macroeconomic equality between income and current output\textsuperscript{40}. Now, '[t]he anomaly [i.e. an excess of demand on the product market] cannot occur any more in a system with the two departments. According to the new rule of the game, each new product is sold necessarily by the expenditure of the corresponding income' (p. 312, our translation)\textsuperscript{41}. In a nutshell, any excess of bank credit (hence any credit-led excess of aggregate demand over current output) is averted by introducing in modern banking the structural bookkeeping distinction between money and credit. By this bi-partition no bank will be able to finance an excess of demand on the market for produced goods and services. The mechanical working of the reformed monetary structure of bank accounts excludes it.

Of course, if there were enough space to expand here, this bi-partition could be developed further, in an attempt to illustrate the whole dynamics of the reformed payment structure and its beneficial implications for the management of credit-led inflationary pressures. However, at this stage of the argument the reader's attention has to be drawn on the analytical principle rather than on its practical implementation in modern banking\textsuperscript{42}. The operational, and clear-cut, distinction between the money-purveying and the credit-purveying department, and the accounting rules governing their relations, are sufficient to avert excessive bank credit to be granted.

\textsuperscript{40} Note in passing that this rise in prices is probably captured (at least in part) by the targeted price index, although analysis of an increase in the latter cannot explain its cause (see Parts I and II).

\textsuperscript{41} 'L'anomalie ne peut plus se produire dans le régime des deux départements. Selon la nouvelle règle du jeu, tout nouveau produit est nécessairement écoulé par la dépense des revenus correspondants' (Schmitt 1984: 312).

\textsuperscript{42} As noted by Schmitt in his "Introduction" to the Italian edition of Ricardo's monetary writings, from a theoretical point of view the reform brought about by the Peel Act in 1844 is less interesting than the underlying analytical principle, i.e. the separation of money's emission from financial intermediation (Schmitt in Ricardo 1985: 82).
On the necessary introduction of a fixed capital department

So far, the Ricardian division of the banks’ bookkeeping enables not only the banking system as a whole but also each bank individually to know exactly the amount of savings deposited by the public. Owing to the structural distinction between monetary and financial operations, no money creation can acquire the pathological status of a creation of credit any more. Thus then, each and every loan provided by any deposit bank originates in production, inasmuch as the sum total lent to banks’ borrowers cannot exceed the exact amount of income deposited by saving units. To say it again, the clear-cut separation of department I from department II mechanically prevents a credit-led excess of demand for available output.

However, such a bookkeeping structure would yet be unable to prevent the birth of empty money in the production process, namely in the investment of the firms’ profit. As a matter of fact, the bi-partition concisely explored in the previous subsection concerns output circulation, and can indeed avoid benign inflation as defined in Chapter 5. It is however on production that our structural monetary reform must ultimately focus, if it aims to get rid of malign inflation. As we have previously noted, the remuneration of wage-earners and the investment of profit in the production of fixed capital are recorded today in the same bank account (that is, department II). This is exactly the principal cause of empty money and, consequently, of the birth of an irreversible inflationary gap in present-day capitalist economies. The bi-partition introduced in the previous subsection is therefore a necessary, but not yet sufficient condition for solving the inflation problem. It resolves in fact benign inflation but cannot avoid malign inflation, which is certainly much more troublesome for the macroeconomy than the former can really be.

Let us recall that the distinction between benign and malign inflation consists in the fact that the former is not cumulative in time, since its effect on monetary equilibrium is destined to be compensated by loan repayment. Malign inflation, by contrast, elicits an irretrievable inflationary disequilibrium between total demand and total supply, as we shall see more clearly in what follows.
As stated by Cencini, ‘[t]he problem within the actual structure of domestic payments is that the formation of fixed capital does not lead to the capitalisation of profits’ (Cencini 1996: 59). It is because invested profits are not withdrawn from ‘financial circulation’ (represented by department II, which is also the sole department in actual bank accounting) that they still form part of the ‘loanable funds’ (Schmitt 1984: 322-323). To repeat the crux of Schmitt’s analysis, conflating the book-entry records for the investment of profit and the payment of the wage bill (in the capital goods sector) amounts to eliciting empty money, because the resulting deposit (recorded in the banks’ financial department) has already been spent when the firms’ profit was invested. ‘Effectively, it is not the investment of profit that has to be questioned, but the way its entry is recorded in bank accounting’ (Cencini 1995: 74).

Let us try to clarify the main line of argument. If the income that firms invest on the labour market were kept within the banks’ financial department, as is the case today, it would define at one and the same time an identical amount of savings at the disposal of any bank’s borrower showing up with good credentials. Within the present structure of bank accounting, in fact, the expenditure of profit on the factor market (productively) consumes the profit, but not the corresponding bank deposits (Schmitt 1995-1996, vol. II: 67-68). This means that the investment of profit in the production process does not simultaneously withdraw an equivalent amount of bank deposits from (financial) circulation, despite the fact that a macroeconomic capital is definitively absorbed in the investment goods newly produced. The income corresponding to the firms’ invested profit can, and will, thus be lent by the banking system to some economic agents, and will be spent in the end on the product market. To this inference one then only needs to add that an income already spent on the factor market (when profit was invested) is spent again on the product market, to conclude that the second expenditure of the same income defines an inflationary gap for its whole amount.
Let us try to avoid a possible misunderstanding. We are not maintaining that profit investment is, in itself, either anomalous or pathological: no monetary disequilibrium can be imputed to the decision of firms to invest their profit on the factor market, neither in theory nor in actual facts. As already noted, it is not the investment of profit in a new production activity that has to be blamed for generating inflationary pressures. It is the present, too simple structure of bank accounting that has to be improved. In particular, as Cencini puts it, 'since the deposits that correspond to invested profits have already been spent in the payment of wages, they ought not to be available to be spent on the product market any more' (Cencini 1999: 143, our translation)\(^4^4\). Today, these deposits are recorded as savings in the banking system, which can therefore lend the corresponding amount for consumption purposes\(^4^5\). This operation elicits an excess of demand on the product market, for it leads to the formation of an income void of any substance\(^4^6\). Since the corresponding goods and services newly produced have already been purchased by firms on the labour market, via the investment of profit, total saleable output is automatically diminished without at the same time identically diminishing the amount of money income nourishing total demand. The result is precisely the definition of an inflationary gap, irremediably confirmed by fixed capital amortisation (as seen above).

Now, since the emission of empty money in the investment of profits on the factor market is due to the direct imputation of the wage bill on firms' profits in the banks' bookkeeping, the solution of inflationary pressures depends on separating the workers' [1]

\(^4^4\) 'I] depositi corrispondenti ai profitti investiti essendo già stati spesi nel pagamento dei salari, non devono più poter essere spesi sul mercato dei prodotti' (Cencini 1999: 143).

\(^4^5\) We leave speculation aside here, since, as we have seen, 'at the end of any purely speculative chain of transactions there always is an output consumption. The case made in the present context is therefore general, despite the simplicity of the argument. See also Chapter 5.

\(^4^6\) To state the obvious, the nominal income thus created takes on per osmosis purchasing power from the existing real incomes, diluting value among an increased number of money units. 'As a consequence of inflation, the content of money, unchanged in real terms, acquires a new numerical expression. A greater quantity of money is needed, therefore, to purchase the same product' (Cencini 1995: 59).
remuneration from the capitalisation of profits as recorded within the banking system. To be more precise, since the pathology of fixed capital amortisation originates in a still imperfect monetary structure of the banks' bookkeeping, the solution consists in reforming the latter so as to make sure that the macroeconomic capital invested in the new instruments of production is definitively withdrawn from the financial market, where at present it can be lent. In simple terms, to be effective, our anti-inflation policy must aim at guaranteeing that the whole amount of invested profits within the national economy does not add to the financial circulation of yet unspent incomes. Indeed, only in this situation would the production of replacement goods elicit a sum of 'full' money in the economy, that is, an income whose purchasing power is real and not merely nominal. Let us show briefly how this can be made to happen in the reformed structure of bank accounting.

Following Schmitt's (1984) analysis, a clear-cut, and operational, separation of profit capitalisation from workers' remuneration requires the institution of a third department in the banks' bookkeeping. In addition to the distinction, first pointed out by Ricardo, between the monetary and the financial department (as we have explored in the preceding subsection), a fixed capital department, III, has to be introduced in the payment structure of our domestic banking systems. The introduction of a department of fixed capital is specifically required to record in it any addition to the instrumental capital (that is, the means of production) of the whole economy, so as to impede the birth of empty money that leads, as we have seen above, to the generation of an inflationary disequilibrium between total demand and total supply. Let us concisely try to study the working of this new department with a simple numerical example, referring as much as we can to those raised earlier on (Table 6.3).
Entries (1) and (2) concern the payment of the wage bill carried out by the bank on behalf of firms 1 and 2. As repeatedly noted earlier on, workers obtain an income (equal to, say, 200 units), formed as a bank deposit. Assuming, to simplify what fundamentally does not affect analysis, that no income redistribution occurs among households, entry (3) records the purchase of firm 1’s output by the set of workers: income holders spend an income of 200 to acquire an output worth 100. By marking up the prices of its goods, firm 1 is able to capture an income of 100. It thus makes a profit corresponding to the financial debt incurred by firm 2 for the remuneration of its workers. Entry (4) is the net result of the first three entries. Now, since the profit gained by firm 1 is not redistributed (e.g. as a dividend to the firm’s shareholders or to the government sector in payment of taxes) but spent for the production of fixed

---

47 For the sake of brevity we omit here department I, which must therefore be considered as implicit in this context, since, let us repeat, any transaction requires the emission of its monetary instrument. Also, the bank here represents the national banking system as a whole.

48 Let us emphasise once again that these entries are recorded in any case, that is, even in the much likely situation when firms have pre-existent deposits. This should become clear as the explanation of this general example proceeds. Impatient readers may be reminded that, analytically, the formation of profit has to be explained before tackling the expenditure of profit on the factor market. Both in theory and practice, the formation of national income always precedes its expenditure logically (though not necessarily chronologically).
capital, it has to be withdrawn from the sum total of savings, because otherwise it will ultimately be spent again on the product market. Entry (5), and its alter ego entry (5'), are the necessary, and sufficient, accounting device to impede that banks can lend the savings which correspond to the sum of profits invested in the means of production (fixed capital). Clearly, firm 1 cannot suffer any loss from these accounting operations, which are purely internal to the banking system. Indeed, at this stage firm 1 still has a bank deposit of 100, recorded as in entry (5'). However, since this is a capitalisation (not a redistribution) of the firm's profit, the corresponding amount must not be available to finance a new wage bill, because otherwise, as we know by now, after compensation in department II it will give rise to empty money and, consequently, to an irreducible excess of demand on the product market. Hence, when a fixed capital department is introduced in the banks' bookkeeping structure of domestic payments, entry (6) is the result of the automatic transfer from department II to III of the sum total of profits capitalised in the current period. This entry between the bank's second and third departments makes sure that no income invested in the production process can still circulate within the financial sector of the economy. By this token, when firm 1 remunerates its workers for a new production of investment or capital goods (supposing a new wage bill equal to 100), the bank records the payment as in entry (7). So, the payment of the new wage bill (entry (7)) is not compensated with the firm's capitalised income (entry (5')) in the bank's accounts.

Since in the reformed structure of payments the remuneration of workers does not conflate with the investment of profit on the labour market, the monetisation of any new production cannot affect the relationship between money and output. No inflationary gap is therefore generated in the production of fixed capital. So much so that wage-earners in the investment goods sector are paid with 'full' money, that is, with an income whose purchasing power is real and not merely nominal. To put it slightly differently, 'since F's debt and credit are not entered in the same department, they do not cancel out: profit no longer finances the payment of wages
and the entire system works without leading to the anomalous emission of empty money’ (Cencini 1996: 59)\textsuperscript{49}. Since the payment of any new wage bill is entered in the banks' financial department, this operation does not affect the sum total of profits withdrawn from financial circulation and recorded within the department of fixed capital.

The final expenditure of a new income (of 100) on the product market confirms that in the new structure of bank accounting the investment of profit on the factor market leaves the relationship between money and output unaffected. The deposit workers in the capital goods sector are paid with (entry (7)), gives them in fact the power to acquire the still unsold stock of wage goods, which are the substance of the forced saving elicited by entry (3). As a matter of fact, entry (8) defines the consumption of the items stocked within firm 2, which thus covers its production costs brought forward from entry (2)\textsuperscript{50}. On the whole, the profit invested by firm 1 is capitalised (entry (5'))\textsuperscript{51} without affecting the same firm's financial debt to the bank's second department for the payment of the new wage bill (entry (9))\textsuperscript{52}. Thanks to the reformed structure of the banks' bookkeeping, the investment of profit on the labour market succeeds therefore in transforming the stock of wage goods (forced saving) into capital goods together with the transformation of the firm's monetary profit into a fixed capital. In this framework, the capitalisation of profits occurs by withdrawing the whole amount of invested profits from financial circulation, that is, from the

\textsuperscript{49} F\textsubscript{1} corresponds to firm 1 in our example. The firm's debt is recorded in entry (7); its credit in entry (5').

\textsuperscript{50} The retail price at which the output of firm 2 is sold does not matter for the present discussion. If these goods (worth 100) are sold at a mark-up over factor costs, the firm makes a profit which, if not redistributed, will be recorded in department III. Analysis in this case will merely replicate the situation of firm 1 in the example considered here, adding no further theoretical issues.

\textsuperscript{51} Were the firm’s profit -- recorded as in entry (5') -- successively redistributed or directly spent by the firm on the product market (instead of being invested on the labour market), the corresponding deposit would be transferred back again to department II, were it would be literally destroyed in the funding of the production costs of the items bought.

\textsuperscript{52} Entry (9) is the net result of entries (6) to (8) in Table 6.3.
banks' second department. Thus then, by averting any emission of empty money, the threefold separation of the banks' bookkeeping guarantees ultimately that any investment on the factor market, even merely for the maintenance of the fixed capital accumulated in the economy, leaves the money-output relationship (i.e. monetary equilibrium) absolutely unaffected and perfectly sound.

Let us conclude this chapter by observing that the achievement of an entirely inflation-proof monetary structure of payments (consistent with the nature of modern money) has a strong link with the need to respect the twofold nature of Real Capital called forth by Sir John Hicks (1974), as the reader interested in the history of economic analysis may have noticed. In particular, the threefold distinction of a monetary, a financial, and a fixed capital department in the banks' bookkeeping could be the clé de voûte for the synthesis between the physical and the financial aspects of the macroeconomic capital existing in any non-inflationary capitalist system:

- Firstly, a clear-cut separation between the first two departments ensures that every deposit recorded in the bank's financial department has a substance, that is, a stock of consumption goods as its objective purchasing power. Impeding any excess of bank credit, this bi-partition makes sure that no deposit can ever be created by a mere stroke of a pen. To exist, any bank deposit has to be produced, since it is the monetary form in which output exists until the latter is sold on the market for produced goods. So much so that the newly produced income (recorded as a deposit in department II) is the monetary measure of the physical stock defining objectively the firm's debt for the payment of the wage bill. This measure,

---

53 If it is capital in the volume sense that is being measured, capital is physical goods; but in the value sense capital is not physical goods. It is a sum of values which may conveniently be described as a Fund. A Fund that may be embodied in physical goods in different ways. There are these two senses of Real Capital which need to be distinguished. I do of course borrow the term Fund from the history, and to the history I now turn. I am going to maintain that the distinction is quite ancient; it divides economists, ancient and modern, into two camps. There are some for whom Real Capital is a Fund -- I shall call them Fundists; and there are some for whom it consists of physical goods. It is tempting to call the latter Realists; but since one wants to emphasize that both concepts are real, this is not satisfactory. I shall venture in this paper to call them Materialists' (Hicks 1974: 309).
occurring on the factor market when the current wage bill is paid, cannot change in
the circulation of the corresponding output (on either the product market or the
financial market), for the bi-partition of departments I and II automatically
prevents excess credit creation.

• Secondly, the introduction of a fixed capital department ensures that any profit
invested in the production process is embodied (to use Hicks’s own word) in the
means of production newly produced. Being an income definitively saved in the
form of capital goods, the sum total of invested profits does not have to be kept in
the financial circulation, for otherwise -- as is the case today -- it would define a
Fund with no saleable object associated with it. The definition of an excess of
demand over total saleable output is precisely given by an inflation of total income
with respect to the volume of goods and services really to be sold. In short, the
inflationary gap is defined by an excess of Fund over an insufficient Matter, to put
it in the phraseology of Hicks (1974).

All in all, the idea that inflation arises when ‘too much money chases too few goods’
is correct, though in need of analytical elaboration. If it is true that excess money
(with respect to existing output) is the cause of inflation, analysis must explain how it
is possible for this disequilibrium really to occur, a question that Friedman (1987: 17)
singled out as the deepest issue in inflation analysis. Once the exogeneity-of-money
view is rejected on logical grounds, as we have tried to do in Part II, advocates of the
endogenous money approach need to elaborate a theory of inflationary disequilibria in
strict adherence to the banking nature of modern money. When this is done along the
analytical lines of the macro-theoretical paradigm adopted in this thesis, it can be seen
that inflation arises out of the emission of what may be labelled empty money, that is,
it originates in a mechanism allowing banks to create a sum of deposits with no real
purchasing power attached to them. When creations of empty deposits occur for the
circulation of output, no irremediable inflationary gap is generated in the economy. It
is when they originate from the monetisation of a production process, namely with the
production of capital goods, that these emissions of empty money elicit a pathological
alteration of the relationship between money and output. The purpose of the structural
reform of bank accounting proposed at the end of this chapter is to make sure that
production of capital goods is financed by savings and not by the emission of empty
money. Assuredly, the investment of profit in the production process is necessary for
economic growth and development. But the way it has been entered up to now in the
banks’ bookkeeping does not yet respect the fundamental requirements of a monetary
system fully complying with the book-entry nature of modern money. It is therefore
towards a structural reform of the banks' bookkeeping that our efforts ought to aim,
to make sure that the daily working of our payment systems cannot alter the
relationship between money and output. In short, according to the analytical
framework informing our macro-theoretical investigation, inflation owes its origin to
a deficient, too simple structure of the bank accounting, and it is thus only a structural
monetary reform that will be able to eradicate the problem.
Aim of the thesis and methodology chosen

The objective of this thesis has been to investigate inflation from a macro-theoretical point of view. This ambitious research programme has taken us in an extended journey into the political economy of inflation, with attention focused on the main conceptual and analytical issues of present-day theory, research, and policy. We surely leave behind a number of points that need further clarification, as well as a list of important questions that remain to be addressed. No attempt has been made to provide definitive answers to all the problems raised. Rather, our efforts have focused on clarifying, and forging ahead, the principal line of argument for an alternative, policy-oriented theory of inflation, integrating it with a conceptual appraisal of conventional wisdom. Time has now come to assess the methodology used to frame our analysis, and to appraise the main results we have been able to obtain. The criticisms raised in the interpretation of the theoretical framework used in this thesis will then be reviewed. We shall conclude this chapter with some perspectives for further research, since we are well aware that our analysis was situated at a very high level of abstraction and that it could just hint at a few issues where a much more refined research work has yet to be done.

A methodological and heuristic assessment of our analysis

Essentially, this study has pointed out the need to undertake a fundamental critical appraisal of traditional inflation analysis, which identifies the latter phenomenon with an ongoing increase in the level of aggregate prices because ‘too much money is chasing too few goods’. The recent renewed concern with measurement problems of aggregate prices both by professional researchers and by a number of government...
statistical agencies, as an aftermath of the so-called Boskin Report\(^1\), has provided us with an apposite starting point to approach the object of the present thesis, namely the origin and the cure of inflation in modern capitalist economies. As noted by Blow and Crawford, ‘measuring “the rate of inflation” -- even defining what we should mean by that term -- remains one of the trickiest and oldest problems in economics’ (Blow and Crawford 1999b: book cover). It is a fact that the conventional method used to measure inflation, to wit, the calculation of the rate of variation of retail (or consumer) price indices, suffers from a number of defects (or biases). Generally speaking, these defects are ascribed to various technical problems affecting the computation of official price indices, where the main sources of biases are (a) commodity and outlet substitution by consumers, (b) introduction of new goods into the marketplace, and (c) quality change in existing goods.

A detailed, yet not exhaustive, survey of the huge literature on the technical issues raised by the measurement of aggregate prices has been provided in the first part of this thesis. Besides attempting to give a bird’s-eye view of the most recent research work carried out in the field of ‘measurement economics’, as Triplett (1975: 19) has called it, in Part I we also aimed to point out, at an intuitive level, that the recorded technical difficulties in measuring inflation by price index analysis might stem, ultimately, from some fundamental problems originating in this very method of inquiry. In simple words, our view of, and contribution to, the ongoing debate on the exact measure of inflation is that the identified measurement problems in aggregate price analysis -- for the calculation of the inflation rate, let us underline it\(^2\) -- represent only the tip of an iceberg whose submerged mass has not been fully

---

\(^1\) See Advisory Commission to Study the Consumer Price Index (1996).

\(^2\) We did not focus on other fields of investigation involved by the mismeasurement of (aggregate) price changes (for instance, standard-of-living evaluation and unwanted distributional issues in indexation processes). Suffice it to say here how enormously relevant these investigations may turn out to be, and how important it is that they should be tackled with alternative, and more sophisticated, analyses than the one attempted in this work.
explored yet. As a gambit to Part II, which focuses on the neoclassical analysis of inflation from a conceptual point of view, Chapter 2 has tried to show that the present difficulties in the measurement of inflation are not merely a matter of technicalities, as current research depicts them, but have an analytical (some would even say ontological) origin. Indeed, the origin of these problems might be twofold: on the one hand, it could be traced back to the quantity theory of money and to the traditional dichotomic view the latter is based upon; on the other hand, it could be ascribed to the still widespread conviction that ‘all of macroeconomic analysis is methodologically and perfectly analogous to microeconomic analysis’ (Boland 1982: 84). As a matter of fact, the methodological approach encapsulated in the currently dominating paradigm of our science is based on the assumption (or research strategy) that there is, and should be, no solution of continuity between micro- and macroeconomics. In this framework, statements that are valid for a single unit are assumed to hold also at the aggregate level or for the economy as a whole. Further, according to this approach, it is reasonable to analyse the economic behaviour of a group of individuals (or of the whole community) by conceiving and modelling the behaviour of a typical or representative agent, and by transferring the results to the aggregate level or to the level of the entire national economy. Surely, the dominant analysis of inflation stems from this view, as we have attempted to show in Part I. Now, the observation that the representative market basket used to compile aggregate price statistics cannot account for the total supply of goods, services and financial assets existing over a national economy, coupled with the parallel analytical observation that, by definition, no sample of consumer units can account for the total demand exerted by income holders in any given period of time, represent, to the best of our knowledge, the original contribution of this thesis to the age-long, and still unsettled, theoretical debate on the so-called microfoundations of macroeconomics (where the multifaceted index number problem is just an epiphenomenon of the specific problem dealt with in this work). As we have pointed out in the last section of Chapter 2, the analysis of inflation by means
of price index calculation seems to us inappropriate to grasp the phenomenon studied. Ultimately, the problem is not technical (as set forth in the literature) but conceptual. It concerns the heuristic status of the explanation of a macroeconomic phenomenon -- the (loss in the) purchasing power of money over total output -- based on what fundamentally is a microeconomic approach (i.e. price index analysis).

Part II has sought to provide further analytical foundations to this line of argument, which incidentally aims to look for macrofoundations of macroeconomics. After an initial discussion of the quantity theory of money in the different formulations this theory has received over the last hundred years, Chapter 3 has tried to develop an internal critique of the dichotomic view upon which this theory is based, a view that considers money and output as two distinct, and (at least in part) autonomous, things. The Cencini (1988) critique of this view, namely that the latter can explain neither the existence nor the value of money in a logically consistent way, has been taken up and forged ahead in the second section of the chapter, where attention has been focused successively on the economic measure of output, on the purchasing power of money, and on the (im)possibility to explain inflation by appealing to the traditional dichotomy between the monetary and the real sectors. Our analysis has drawn attention to three fundamental, and closely related, problems: (a) output cannot be measured in strictly economic terms unless it is associated with money from the very moment of its formation on the factor market; (b) the purchasing power of money cannot be determined objectively if money is introduced on the commodity market, where it is put side by side with a heterogeneous aggregate of physical output; (c) nothing can be said on inflation in a framework where output and money remain alien from one another and where the money stock has been increased by some invisible hand, since in this case only an increase in the price level can logically be inferred from the money injection, with no further insight into the alteration of money’s

---

purchasing power (because of point (b)). If our analysis is correct, these are three main problems which neoclassical monetary macroeconomics will need to address in the future. Present elaborations of this framework hardly go beyond the attempt to introduce money into general equilibrium models where goods already exist in the personal endowments of the contracting parties\(^4\), while the very economic process which associates money and (newly produced) output remains generally unexplored on account of the dichotomic approach.

An attempt to remove the dichotomy between the monetary sector and the real sector from current macroeconomic analysis, and to indicate a potential avenue for the exploration of the association between money and production, has been provided in Chapter 4, where emphasis has been laid on the nature of modern money within a monetary economy of production. In an attempt to introduce a conceptual distinction between the means and the object of payments grounded on the Smithian distinction between money proper and money's worth, this chapter has shown that viewing money as the general equivalent of non-money goods raises the same fundamental problems as those raised by the dichotomy between the real and the monetary sectors of the economy. Leaving aside at this stage the picturesque idea of helicopter money (or of money trees), the discussion has focused on Cartalism, i.e. the theory that money is the only legally valid means of payment, for it alone has the \textit{imprimatur} of the State (which may claim it back in payment of taxes)\(^5\). In this respect, the discussion in Chapter 4 has drawn attention to two crucial points, which deserve

\(^4\) Part II has critically considered cash-in-advance models and overlapping-generations models of money-using economies, as two recent frameworks that have been put forth to account for both the existence and the value of money. An analogous critical assessment could have been carried out about money-in-the-utility-function (or in-the-production-function) models.

\(^5\) As we have underlined in Chapter 4, if money were ‘[a] fiat currency produced monopolistically by the government’ (Fama 1980: 50), it would logically pertain to the category of goods and services defining national output. Yet, including money in the set of commodities would raise the problem of measuring goods by means of goods, a problem that Ricardo had been trying to solve without success until his death. As explained by Cencini (1988), in fact, Ricardo's search for an invariable standard of value was bound to fail, since only a dimensionless standard (i.e. mere numbers, with no dimension whatsoever) does not vary.
further elaboration: (a) no economic agent, even as powerful as a government may be, can really pay its purchases of goods, services (including labour services) and financial assets by simply letting its own IOUs circulate within the national economy; (b) if this exchange of State-issued money against real goods and services occurred in reality, no objective measure of the value of the exchanged items would possibly exist, and the purchasing power of (fiat) money would be indeterminate, too. As a matter of fact, while the State can determine by a political decision the legal tender, or ‘that which is necessary to pay taxes’ (Wray 1998b: 4), it cannot determine its purchasing power, since the latter results from economic activity. In an attempt to elaborate on this latter point, a very simple, and highly stylised, model of the monetary circuit has been presented in the last subsection of Chapter 4. Stemming from the macro-monetary analysis developed over the last fifty years by the Dijon-Fribourg School (led by Bernard Schmitt and Alvaro Cencini), this model has pointed out two features which might be of interest for the construction of the complete theory of a monetary production economy within the endogenous money approach: (a) money and output are not two distinct, and ontologically independent, things, but the numerical form and, respectively, the real content of any monetary transaction; (b) any payment on the factor market, on the financial market, and on the market for produced goods and services concerns a single object, which is literally transformed by the exchange between the monetary flow and the real flow occurring with the intermediation of a book-entry system of accounts provided by banks.

Since ‘the Schmitt-Cencini strategy is to adhere rigorously to the discipline of double entry bookkeeping’ (Desai 1988: xii), the framework developed in this thesis

---

6 It is a fact that no-one can pay by issuing his or her own acknowledgement of debt. This is why money is issued in a triangular operation, involving a bank, a payer, and a payee: the payer needs a third-party acknowledgement of debt to really pay his or her due. See Schmitt (1966) for analytical elaboration, and Rossi (1998: 44-50) for an analysis of interbank payments along the same lines.
has given prominence to a macroeconomic analysis based on banks' balance sheets (owing to the banking nature of modern money). In fact, as argued by Graziani,

[...] this framework seems to me to be relevant for at least three reasons. The first is that it enables to analyse the economic process as a directly monetary phenomenon [...] . The second is that it clarifies unequivocally that an authentic monetary economy uses only credit-money, for any commodity-money, being in reality a commodity, cannot dispose of barter trade. The third is that it enables to see the economic mechanism as a problem of social relations, in a world where the group of capitalists-entrepreneurs is set against the group of wage-earners (Graziani 1985: 131-132, our translation).

The simple model of the monetary theory of production developed here is a macroeconomic model, concerned with output as a whole and where any monetary transaction involves three poles: the payer, the payee, and the banking system as an intermediary. It is very far removed from the forms of economic behaviour of individual agents, who decide on their own level of production, consumption, investment, and so on. Our macro-monetary analysis has proceeded indeed at a high level of abstraction, with arguments developed within the framework of very stylised cases. The exclusion of international trade, exchange rates, interest rates, governmental activities, economic growth, and financial crises are surely important omissions, and the model used to frame our investigation can best be viewed as a basic one, to which these (and many other) features should be added, as we shall point out later on.

It is within such a simple (yet not simplistic) framework that in the last part of this thesis we have tried to study the causes of inflation, defined as a macroeconomic disequilibrium between total demand and total supply of current output. Chapter 5 has

---

7 'Questa impostazione mi pare rilevante almeno per tre motivi. Il primo è che essa consente di analizzare il processo economico come fenomeno direttamente monetario [...] . La seconda è che essa chiarisce in modo inequivocabile che una autentica economia monetaria fa uso unicamente di moneta creditizia, in quanto ogni moneta merce, proprio in quanto merce, non può non restaurare surrettiziamente il regime del baratto. La terza è che essa consente di vedere il meccanismo economico come problema di rapporti sociali, in un mondo nel quale il gruppo dei capitalisti-imprenditori si contrappone a quello dei lavoratori salariati' (Graziani 1985: 131-132).
been concerned with two main topics: the macroeconomics of the wage bargain and the role of bank credit policies in respect of inflation. At the outset of the chapter our analysis has focused on an idea which has a wide currency both in academic quarters and in public opinion, namely that inflation originates in the discrepancy between the growth rate of money wages and the growth rate of (average) labour productivity. Though the results we have been able to reach at this stage represent little more than a few macroeconomic elements on which further research needs to be done, they have nevertheless highlighted a question of fundamental importance: is it ontologically possible to measure labour productivity in economic terms independently of the money wages distributed in remuneration to the workers whose productivity is the object of measurement? The analytical elements put to the fore in the first section of Chapter 5 have pointed to a negative answer. To put it picturesquely, they have shown that labour productivity and money wages are like the volume of a gas and the volume of the room in which the gas is released. In economic analysis, labour productivity cannot be measured if it is abstracted from the workers' remuneration, because the material result of their activity is an incommensurable collection of miscellaneous things having a multitude of different physical attributes. It is thus necessary to resort to money in order to obtain a homogeneous measure of national output, because money units are homogeneous at any point in time. Ultimately, such an approach has led us to understand that the wage bill paid for the production of any good is the exact measure of the latter, since the former is precisely defined by the association of a sum of money with a 'real thing' produced by labour.

In order to explore further the monetary macroeconomics of the labour market and its link with inflationary pressures, in the first section of Chapter 5 we then have briefly investigated the well-known conflict theory of inflation, i.e. the idea that

---

8 Physical heterogeneity also concerns the physical measure of labour (i.e. the measure of output in terms of labour-time). As we have shown in Chapter 5, only the monetary measure of labour, i.e. the workers' remuneration, can dispose of this problem.
inflation arises out of the class struggle for income shares (basically, the real wage and profit shares). Focusing exclusively on the wage-price spiral generated by the struggle for the distribution of national income between different functional categories of economic agents, seems not to be sufficient to provide a theory of inflationary disequilibria up to the task. As a matter of fact, since the distributional conflict between, and within, the various groups of agents cannot modify the existing amount of income, the macroeconomic relationship between money and current output can never be modified by this conflict. It is certainly possible that the real wage and profit shares claimed, respectively, by workers and capitalists are inconsistent at the aggregate level, and that the functional distribution of current total income is the contingent outcome of the relative strength of each class on the labour market as well as on the product market. However, this class struggle can generate no inflationary disequilibrium in the money-output relationship, that is, in the relationship that defines the purchasing power of money — although it may of course elicit an increase in the aggregate price level, which is a further proof of the need to keep inflation analytically separate from changes in the targeted price index.

In an attempt to seize the mechanisms leading to the formation of an inflationary gap, that is, a disequilibrium between total demand (i.e. national income) and total supply (national output), the second section of Chapter 5 has been concerned with the macroeconomic consequences of the credit policies followed by commercial banks for the implementation of their sales- or profit-maximisation strategies. As Post-Keynesian monetary economists have pointed out⁹, the credit facilities provided by commercial banks to the economy may be granted beyond the level that would ensure monetary equilibrium. This means that the relationship between money and current output may be affected negatively by an excess of bank credit, which would thus elicit an inflationary rise in goods and assets prices in actual markets. This is a serious issue

---
that will need further attention, particularly at a time when speculation activities in financial markets represent an increasing share of total transactions, as clearly illustrated by Howells (1999), and Howells and Hussein (1999). Our own contribution to this important line of research has tried to show that excess credit facilities are not irrevocably inflationary for the economy as a whole, for their effect on the relationship between money and output is not cumulative in time. In fact, bank loans have eventually to be repaid by the banks’ borrowers. Hence, the inflationary gap elicited by an excess of bank credit granted over any particular period of time is destined to be compensated by a disequilibrium of the opposite algebraic sign when these very bank loans are reimbursed, that is, when bank deposits are reduced by loan repayment\(^{10}\). Following the terminology first used by Schmitt (1984), we have labelled this credit-induced mechanism as a ‘benign inflation’, to point out that it is not definitively harmful for overall monetary equilibrium. To illustrate the benign nature of this phenomenon at the macroeconomic level, reference has been made to personal sector borrowing for second-hand (or purely speculative) transactions. Attentively considering the nature of credit-induced bank deposits with only a remote link to production has indeed provided us with the opportunity to slightly expand the Schmitt framework, since we have applied the latter to a problem yet unexplored by this School of thought. As some critics have in fact pointed out (Desai 1988: xiii-xiv; Deprez 1989: 204), the Schmitt-Cencini analysis has well investigated the link between money and production but has neglected so far to explore the realm of financial speculation. Our discussion of excess credit facilities has indicated the broad lines along which this omission could be extenuated, as far as the inflation problem is

\[
\text{\textsuperscript{10} Of course, this does not exclude that a new inflationary gap may arise out of the new loans granted by the banking system at precisely the same time when the previously granted loans are reimbursed. It is in fact this continuum of newly granted bank loans that makes it 'empirically' impossible to observe the compensation we drew attention to when analysing the theory of excess credit facilities.}
\]
concerned. Needless to say, other, and more important, contributions are to be made in the field of speculation, where research has just begun\textsuperscript{11}.

The last chapter of this thesis has sought to push forward the Schmitt-Cencini analysis on another front, which proves to be much complex and difficult to deal with. Indeed, to the best of our knowledge the link between inflation and the monetary macroeconomics of fixed capital replacement has not yet been investigated by our profession. Chapter 6 has aimed to point out a few macro-analytical elements that could help to fill this gap in current research, although we are well aware that we could not take the matter at any depth in the scope of this dissertation -- whose principal objective, let us recall it, was to provide a fundamental critique of traditional inflation analysis, grounded as it is on the concept of the price level and on the measurement of changes in the latter. In the first section of the chapter we set out to polish up the analysis of what the Dijon-Fribourg School defines as an emission of 'empty money', linking the problem of inflation to the process of capital accumulation going on in the macroeconomy. In particular, by focusing attention on the amortisation of the fixed capital accumulated in the entire national economy, we have been able to point out that national income determination must take into account the value of the replacement goods currently produced, since the latter production elicits a wage bill which logically enters into the measurement of $Y$. This question has been unnoticed so far in economic analysis, Schmitt's (1996b) theoretical analysis of unemployment notwithstanding. It might however be an important issue to consider for the construction of a modern, and policy-oriented, theory of inflation, because it might represent a crucial element in order to explain why an excess of demand on the market for produced goods and services may occur in reality.

\textsuperscript{11} As Chick has pointed out, 'economists, at least until very recently, have resisted developing a theory of speculation and incorporating it into the main body of economic theory' (Chick 1994: 380). See note 9 for references to leading research on this important topic.
Granting the importance of this issue, on which a research programme aimed at elaborating the analysis sketched here is called for, let us try to reconsider it very briefly. If it is correct to claim that \( Y = C + I + U \), where \( U \) is the total value of replacement goods, it then follows that total demand \( (Y) \) is greater than total supply since the latter consists of final goods only (that is, \( C + I \)). As a matter of fact, replacement goods are intermediate goods, because they replace the fraction of fixed capital used up in the period; as such, they are not for sale on the product market, where only consumption and investment goods are supplied. Thus, in spite of the fact that total supply and total demand are identically equivalent as measured by the total wage bill paid in the current period, total demand is also greater than total supply in terms of available output. Things would not be problematic at all, in so far as the inflation problem is concerned, if replacement goods were indirectly bought by deposit holders on the market for consumption goods, where retail prices may include a mark-up factor that accounts for the production costs of amortisation goods. However (but this point needs further research), things change radically when, as in the present structure of banks' bookkeeping, the investment of the firm's profit in a new production process (i.e. capital accumulation) elicits a sum of bank deposits deprived of purchasing power over current output. As we have argued at length in Chapter 6, it is in fact the payment of the current wage bill out of pre-existent deposits that at present opens up the way for the birth of an irreversible inflationary gap in our monetary economies of production. By paying their wage bills out of previously accumulated profits, firms purchase on the factor market the very output wage-earners are remunerated for, so that the bank deposits the latter are credited with are emptied of any goods newly produced.

The problem we drew attention to in the last part of this work has nothing to do with the forms of behaviour of decision-making agents. As a matter of fact, the investment of profit is sound and necessary for economic development. But the way it has been entered up to now in the banks' bookkeeping does not seem to respect the
book-entry nature of modern money and the fundamental distinction that exists between money, income and capital. This line of reasoning has led us to the conclusion that in this framework the origin of inflation is 'monetary' and 'structural' rather than 'real' and 'behavioural'. The policy proposals we have elaborated in the last section of Chapter 6 have therefore aimed at improving the structural monetary framework of modern banking and payment systems. In particular, our analysis has focused on refining and pushing forward the structural monetary reform called for by Cencini (1995: 75-76; 1996: 58-59), who -- following the principle put to the fore by Schmitt (1984: 322-324) -- made an attempt towards spreading monetary transactions over three functionally distinct bank departments, in order to have an inflation-proof economy. On this point, the contribution of this thesis is twofold. With the help of a simple, and very stylised, numerical example, the distinction between money and credit (that is, the distinction between monetary emission and financial intermediation) has been explained in bookkeeping terms, with emphasis laid on second-hand (i.e. non-productive) transactions in housing assets, so as to provide an explicit link with the discussion of excess credit facilities in Chapter 5. Further, the introduction of a fixed capital department in the banks' bookkeeping has been explained in detail, by means of a two-period example of production, consumption, and capital accumulation that has separated, in book-entry terms, the investment of profit in a new production process from the payment of the corresponding wage bill.

Clearly, the reform called for at the end of this thesis needs further elaboration to be more convincing and to allow for its heuristic assessment. Our outline of the structural changes of the banking system required to make sure that no inflationary disequilibrium can occur is certainly very rudimentary and not sufficient to provide a valid substitute to traditional inflation analysis. Yet, given the object of this thesis, it has not been possible to pursue this matter any further here. Its elaboration requires much more work that could have been done in this context, and that can presumably be demanded from a single researcher, too. Surely, the entire book-entry skeleton that
could only be drafted in the second section of the last chapter needs re-examining. A number of questions will emerge from the exploration and discussion of the proposed new monetary structure of bank accounts that has been undertaken at the end of this work. Indeed, a reader's first reaction might be sceptical about both the usefulness and the extent of such a plan of reform, especially since in practice the latter amounts to nothing more than changing bookkeeping rules. It might even be the case that the proposed distinction of the three departments (money, credit, and fixed capital) in the banks' bookkeeping will prove to be impossible to realise, as has been claimed by Vallageas (1988: 189) in connection with the proposition to separate the monetary department from the financial department\textsuperscript{12}.

The first, natural line of research following upon this work should aim precisely at further investigating the structural distinction between the three bank departments, to establish whether or not it is really feasible, and to appraise its normative content. On this last point the greatest caution is called for. It is in fact at this level that a number of criticisms may enter the scene, as they have done in the (very rare) literature concerned with the analytical framework used in this thesis. To our knowledge, no attempt at providing a taxonomy of these criticisms has ever been made up to now by either critics or supporters of this framework. We offer it here, since it might also be helpful to indicate a number of lines of research in order to expand, and also to assess, the paradigm adopted to approach the specific problem investigated in this thesis.

\textsuperscript{12} As we have noted in Chapter 6, Vallageas (1988) claims that this distinction is impossible on account of the variety of bank loans. '[The monetary and the financial departments] could be distinguished if their loans would have different forms. [\ldots] In reality, banks lend against a great variety of claims’ (Vallageas 1988: 189). '[Ces départements pourraient être différenciés si leurs prêts revêtaient des formes différentes. [\ldots] Dans la réalité les banques accordent des prêts contre des titres extrêmement variés’ (Vallageas 1988: 189).] This statement does not seem to consider that the emission of money must not be confused with the granting of a loan by the bank, as we have pointed out in Chapters 4 and 6. See Rossi (1998, 1999) for more details.
Criticisms raised in the interpretation of Schmitt's theory

On the definition of money

First of all, there is the basic question of the new analytical conception of money, defined as a numerical form, or as an immaterial vehicle, of which physical output is its real content, or load. The first (to our knowledge) published critical review of the Dijon-Fribourg School -- by a supporter outside it -- noted that in this framework 'money is distinct from its purchasing power “as blood is distinct from its oxygen”' (Devillebichot 1969: 693, our translation). More recently, a reviewer of Cencini's 1995 book positively stated that the Schmitt-Cencini theory offers 'powerful insights into (theoretical) monetary issues via an analysis of money in a world of banking, where money is essentially a book entry' (Cohen 1996: 1134). As was expected by Lord Desai in his "Foreword" to Cencini (1984), this analysis has 'puzzle[d], stimulate[d], infuriate[d], or annoy[ed] many readers' (Desai 1984: xi). Essentially, these reactions can be ascribed to the transcendent definition of money put forward by the Dijon-Fribourg School, that is to say, to an analytical conception of money transcending the agents' perception of the instrument they use to accomplish their monetary transactions. In particular, Schmitt's theory of money appears to some critics as being 'nothing more than the fact that when a bank creates money it simultaneously creates an asset and a liability' (Watkins 1985: 596). Certainly the asset-liability nature of bank money challenges a number of monetary analyses. So much so that it has elicited an analysis in terms of circular flows where the numerical form of any monetary transaction is created and destroyed in the same instant, that is,

---

13 'La monnaie est distinctive de son pouvoir d'achat “comme le sang de son oxygène”' (Devillebichot 1969: 693). This analogy is intended to point out that blood (money) is the vehicle of oxygen (money's worth, that is, physical output). Notice here the similarity with the idea of 'the great wheel of [output] circulation' put to the fore by Smith (1776/1970).

14 See Chick (2000: 130), who complains that the Schmitt-Cencini definition of money does not conform to the statistical definition used by central banks and international organisations.

258
when a payment is made. This has puzzled those critics who question 'why the value of something (or some action) which is, by [the Schmitt School] description, evanescent, or destroyed almost as soon as it is created, should be of interest' (Chick 2000: 129). An attempt at clarification has been provided at the end of Chapter 4. In fact, a subsidiary aim of this dissertation has been to clarify (of course, within the limited scope and boundaries of the present work) the theoretical approach used in it, couching the Schmitt-Cencini framework in simple, or more familiar, language. Surely, further intellectual efforts are required to build a bridge with other Schools of monetary thought where this possibility might exist, although this bridge-building cannot come from an individual's effort alone.

On the relationship between money and production

Secondly, and closely related to the preceding point, there is the fundamental link between money and production, which may require further elaboration. In his book review of Cencini (1988), Deprez has claimed that '[t]he [Cencini] discussion of the emission and circuits of money and income is based around a classical conception of the period of production, which is implied to be uniform and synchronised for all production in the economy' (Deprez 1989: 203). A similar point had in fact already been raised by Graziani (1985: 139), who questioned on practical grounds the idea that output is saleable at the precise instant when wages are paid out. There seems to exist here a problem in the case of those goods (say, a Boeing 777) which require several production periods -- as defined by the periodic payment of wages -- in order to be completed and put on sale on the product market. An analogous, and perhaps more challenging, problem seems also to exist with those goods and services which

---

15 See Chapter 4.
16 This thesis has made a first step towards bridge-building in respect of the problem of excess bank credit and its link with inflationary pressures (see Part III).
17 See Chapter 5 (note 20).
cannot be sold directly to the holders of bank deposits (recall Keynes's famous example of wage-earners digging holes in the ground\textsuperscript{18}), and with those goods which have been produced but for which there actually is no market demand (for instance, those out-of-fashion items piled up in the firms because no consumer wants to buy them). Given the importance of these issues, let us very briefly indicate how they have been addressed by the Dijon-Fribourg School\textsuperscript{19}.

The case of those goods and services which cannot be sold directly to income holders may be illustrated referring to public sector production. It is indeed a well-known fact that many public goods and services are not sold directly to income holders, who cannot buy (say) the protective services of the armed forces as they buy a chocolate bar at a shop. The nature of public sector output is such that, as a general rule, the general government sector cannot sell its products on the market. Therefore, being unable to cover all its production costs in the usual manner (on the product market), the State uses taxation for the ‘selling’ of its output. In fact, as far as the relation between money and production is concerned, production by the general government sector does not differ from private sector production\textsuperscript{20}. ‘Whether it is a matter of weapons, welfare services or mere holes, public enterprise output is an integral part of national product and plays a part in the determination of global supply in the same way as wages paid out by public firms do in the determination of global demand’ (Cencini 1995: 65). This does not mean, of course, that the physical characteristics of the general government output is of no importance. For both individual and social welfare it is surely more important to produce (say) national health services than nuclear weapons. Needless to say, similar alternatives exist also

\textsuperscript{18} See Keynes (1936: 220).

\textsuperscript{19} See particularly Schmitt (1984: 94-123).

\textsuperscript{20} Chick has recently argued that the Schmitt School puts to the fore a ‘production theory of money’ (in contrast to Keynes’s ‘monetary theory of production’), because Schmitt’s theory is founded on ‘the generative power of labour to create income and the role of this activity to give value (purchasing power) to “money”’ (Chick 2000: 129).
in the private sector economy. This important subject matter lies however far beyond the scope of our work.

The case of unsaleable goods can be dealt with in a similar way. If a firm has produced a stock of unwanted goods and cannot find a single consumer wanting to buy them, it encounters a loss (or a decrease in its profit) since it cannot cover the corresponding production costs on the product market. The problem can be analysed following two alternative approaches. If one measures income at the time of sale, the existence of unwanted goods piled up in a firm implies that this very firm has to buy them, for it must cover their costs of production. This is done by a decrease in this firm's profit, a decrease which may be of such an amount to give rise to a net loss in the firm's cash flow statement. The firm's future production plans are thus much likely to be revised -- and perhaps to such an extent that the firm will cut back on the employment it offers --, in order to avert the further heaping-up of unwanted goods. The same conclusions are reached if one defines income at the time of production (i.e. when the wage bill is paid), although this approach is different from the previous one\textsuperscript{21}. In this framework, income exists in the form of a bank deposit as soon as wages are paid out to workers. Even if these workers' output is destined to be unsold to income holders, its production generates the exact amount of income necessary and sufficient to cover the corresponding factor costs. Now, since income is in the form of a bank deposit, banks will lend the corresponding amount which is not spent by income holders. This implies that those goods which are not bought by consumers are bought by the firm itself, because, let us repeat it, firms must sell unwanted output to cover its costs of production. By lending to firms, banks therefore guarantee the identity between total demand and total supply of current output at any point in time -- which is another proof of the necessity to search for the causes of inflation beyond the analysis of consumers' or producers' forms of behaviour (see Part III).

\textsuperscript{21} Besides measuring income at market prices, national accountants also measure income at factor costs.
A third set of critical remarks concern the fact that the Schmitt-Cencini view ‘results in the production periods being isolated from each other’ (Deprez 1989: 204), an isolation that accounts for the unitary value of the multiplier. This probably is the most recurrent critique raised in the literature published up to now. It has been pointed out twice by Lord Desai (1984: xi-xiv; 1988: xiii), and results from the ‘refusal [by the Dijon-Fribourg School] to recognise that income earned in one period is spent in the next and, through being spent, generates new income in a still later period’ (Hennings 1985: 621). To put it differently, this criticism amounts to claiming that ‘the role of investment goods, the saving propensity of workers, and the consumption of capitalists or rentiers is not explored [by the Schmitt-Cencini analysis]’ (Deprez 1989: 204). In this framework, according to Chick, ‘investment is not singled out for special treatment: it does not matter whether production takes place in the consumption goods industries or the capital goods industries’ (Chick 2000: 129). It is in this connection that Lord Desai pointed out that ‘the Dijon School’s theory is still an equilibrium one and leaves one very little room for disequilibrium dynamics’ (Desai 1988: xiii). As he had already stated back in 1984, ‘[t]he multiplier process is a disequilibrium dynamic one whereby output changes from one level to another’ (Desai 1984: xii). This amounts to saying that the Dijon-Fribourg School has no disequilibrium dynamic theory.

As a perusal of Schmitt (1984: 39-74) will show, the Schmitt theory is however neither static nor dynamic, because it is ‘quantic’ (see Realfonzo (1999: 377)). The idea of ‘quantum time’ is indeed at the core of this theory (see e.g. Schmitt (1982)), for which any production process is an emission of a quantum of time (that is, a finite and indivisible period of time taken as a whole) as soon as it is monetised by banks.

---

22 In his recent account of the French circuit School, Deleplace (1999: 475) notes that neither production nor income distribution creates a link between periods, so that the multiplier is necessarily equal to one.
'We have a quantum theory of production because the act of producing is a wave, which quantizes production time' (Schmitt 1986: 127). The production period is precisely defined by the periodic payment of wages. For example, if the latter are paid on a fortnightly basis, the former has a duration of two weeks (fifteen days); if the latter are paid once a month, the former has a duration of four weeks (thirty days), and so on, independently of the number of working days (or hours) necessary for manufacturing output physically (Schmitt 1984: 437-448). In this framework, production and consumption are two actions whose result is, respectively, the creation and the destruction of current income via two emissions of the same quantum of time (e.g. two weeks). According to this theory, '[t]he final expenditure of income is, in fact, an emission, and therefore as such it also defines a finite and indivisible period of time. Moreover, creation and destruction being related to the same income, this second emission defines the same quantum of time defined by the first emission' (Cencini 1984: 191). In this framework there is therefore no functional link between production periods, because expenditures are defined in quantum time and depict a circular flow representing a production-consumption (that is, a creation-destruction) of the same money income (pp. 94-162). Hence, there is no 'earning through spending' causal mechanism, in so far as expenditures on the market for produced goods and services cover the production costs of sold output. There is nevertheless a sort of link between periods, since a firm's loss in, say, period 1 (due to a failure to sell all its output on the product market -- see above) is much likely to lead this firm to reduce its level of activity (hence to cut back on the employment it offers) in the following period(s), according to the principle of effective demand explored in Chapter 5.

On economic crises
An important corollary of the latter critique has raised concern with the exclusion of economic cycles from the Schmitt-Cencini view. This has led a few critics to claim
that this theory applies to a Say's Law (real-wage) economy\textsuperscript{23}, and that it maintains that crises are outside the field of economic analysis (Lavoie 1987: 87)\textsuperscript{24}. A slightly different version of this criticism maintains in fact that the monetary analysis of the Dijon-Fribourg School has no place for unemployment (dis)equilibria, so that it would be a theory of a dream-world or, at least, an explanation of an ideal (utopian?) state of the world. Certainly the unemployment problem cannot be neglected by any economic theory whose adherents claim to have something to say about the real world. Yet, it would also be unfair to address this critique to the Schmitt-Cencini view, as a perusal of Schmitt (1984, 1996b) and Cencini (1995, 1996) will show, although nobody denies the need for further contributions in order to construct a fully-articulated, and policy-oriented, theory of unemployment within that framework.

Some perspectives for further research

It is particularly on the relationship between inflation and unemployment that future research seems to be much needed, along the lines informing the macro-theoretical analysis attempted in this thesis. In our highly capitalised economies the accumulation

\textsuperscript{23} On the distinction between a real-wage economy and a monetary economy of production, particularly in connection with Say's Law, see Rotheim (1981: 575-584). Now, the monetary interpretation of Say's Law shows that the latter is in fact a sales-purchases identity (see e.g. Schmitt (1975)): for every economic agent the total (outgoing) flow of expenditures is always identically equal to the total (incoming) flow of earnings over any given period of time, because there can never exist either a net purchase or a net sale for any agent in the whole economy. An agent's net purchase on the market for produced goods is inevitably funded by a net sale of financial assets (for the same agent), because any cession of a bank deposit -- either owned or borrowed -- in exchange for a commodity defines the cession of a financial claim over domestic output. See Rossi (1998: 34, n. 27) for analytical elaboration.

\textsuperscript{24} In connection with the explanation of 'crises' by the Dijon-Fribourg School, Realfonzo has noted that 'Schmitt puts forward a new interpretation of Say's Law' (Realfonzo 1999: 377). It is a pity that Realfonzo could not expand on this crucial point his entry on the French circuit School in the recently published Encyclopedia of Political Economy. Had he done it, he would have noted that it is possible to observe a 'Say's Law economy' affected by unemployment. Recall the previous explanation of unsaleable goods on the product market: those goods not bought by consumers are bought by producers, for the latter must cover their costs of production. Thus then, although the consumers' aggregate demand can differ from aggregate supply, total demand is always identical to total supply.
of fixed capital may reach such an extent that its remuneration becomes more and more difficult. On account of the decrease in the rate of profit in respect of the interest rate paid on the financial market, firms may decide to stop investing in the production of capital goods and to invest their profits for the production of consumption goods. ‘By so doing, however, they would provoke an increase in the supply of these goods without matching it with an equivalent increase in demand. The subsequent deflation would force them to reduce their activity, and, with it, the level of employment’ (Cencini 1995: 90). No-one doubts that this is an extremely important topic on which further research is needed. As Cencini and Baranzini have in fact pointed out in their "Introduction" to a recent co-edited book, where an international team of distinguished contributors investigate inflation and unemployment from different analytical points of view, ‘the coexistence of two apparently complementary disequilibria such as inflation and deflation is still a mystery, and very little progress has been made towards explaining the nature of their relationship’ (Cencini and Baranzini 1996: 1). In this connection, Schmitt (1984: 224-239) has argued that in our advanced capitalist economies \( \frac{1}{3} \) of total current output is appropriated by the set of ‘disembodied’ firms, an appropriation which is the cause of those macroeconomic disequilibria such as inflation and unemployment. This question needs to be further explored, in order to determine whether this appropriation can really attain this proportion (\( \frac{1}{3} \) of total production) and, in the affirmative, why this is so. Furthermore, apart from the important omissions previously noted (especially the role of the general government sector, the interest rate problem, and international trade\(^{25}\)), one should also mention at least three other main topics for further inquiry, where much research has to be done within the theoretical framework used in this thesis.

Firstly, as we have already mentioned earlier on, the problems of financial speculation and of endogenously generated financial crises need a great deal of attention. ‘Indeed many of the trillions of dollars that were wiped out on 19 October, 1987 were fictitious capital to which nothing “real” corresponded’ (Desai 1988: xiii-xiv). Futures, derivatives, forward contracts, credit swaps, exotic options, and an ever increasing number of other financial engineering operations are all topics that the Dijon-Fribourg School has the ability to say things about, but up to now no attempt has been made to investigate them. The simple model used in this work could be extended to incorporate all of them, so as to gain further insights into the distinction between benign and malign inflation discussed in Part III (particularly with respect to the financial circulation of credit-induced bank deposits).

A second element concerns the functional distribution of income within the Schmitt-Cencini framework. The explanation of the formation, and expenditure, of the firms’ profits within the macroeconomic circuit of money wages given in Chapter 5 (see also Schmitt (1966: 283-297; 1984: 123-150), and Sadigh (1988: 47-54)) might integrate some considerations on the ‘optimality’ -- whatever one wants to define by that term -- of income distribution between the real wage and profit shares, perhaps in respect of economic development or growth theories. The purpose of these extensions would be that of providing a more elaborate framework where the inflationary disequilibria investigated in this dissertation could be analysed within alternative models of our monetary economies of production. Also, in connection with the first possible extension, financial transactions ‘build up an intricate network of interpersonal, and private–public, relations, which go beyond those emerging from production activity, and which may profoundly affect through time the distribution of income among the members of the economic system’ (Pasinetti 1993: 86). These issues have an important impact on the process of capital accumulation and are worth exploring.
Thirdly, one might take up the problem of income distribution and capital accumulation in connection with the process of structural change and economic dynamics pointed out by Pasinetti (1981, 1993), to investigate if the different mark-ups across economic sectors or between different industries affect profit determination in such a way as to bear an influence on the inflationary pressures elicited by capital replacement within the macroeconomy. It is clear that the modification of the profit share in the aggregate may have an influence on the extent to which the production of investment goods elicits a sum of 'empty money'. It is less clear, at present, how different mark-up rates may affect the proportion of fixed capital replacement within the total gross investment of firms. A research programme in monetary analysis might be conceived to address these important issues and to expand the framework used in this thesis.

Epilogue

Before bringing this chapter to an end it may be worthwhile to say a last word on our macro-theoretical analysis of the inflation problem. The principal objective of this investigation has been to provide a fundamental critical appraisal of received inflation analysis, centred as it is on the idea of the price level and on measurement of changes in the latter. As long as the problem of inflation is tackled within the conventional framework dichotomising the economy into a real and a monetary sector, no policy-oriented theory of inflation can be elaborated with the certainty that it is not ill-grounded in economic principles. As we have indicated in the last part of this dissertation, a renewed macro-theoretic analysis of inflationary pressures may therefore be particularly relevant for elaborating, and implementing, an anti-inflation policy that could be successful, although many academics and economic policy makers might consider this research programme too challenging to embark upon. We hope nevertheless that our modest attempt at contributing to solving the inflation
problem will stimulate the intellectual curiosity of the reader and lead to further developments of this research programme.
BIBLIOGRAPHY


Cencini, A. (1999) "Inflazione", in *Capitoli di teoria monetaria*, ("Research Laboratory of Monetary Economics", 1), Lugano and Bellinzona: Centre for Banking Studies and Meta-Edizioni, pp. 113-147 (Chapter 4).


Michael, R. T. (1979) "Variation across Households in the Rate of Inflation", *Journal of Money, Credit, and Banking*, vol. 11, no. 1, February, pp. 32-46.


Rossi, S. and Dafflon, B. (1996) "La logique des critères budgétaires du Traité sur l'Union Européenne: premiers éléments d'analyse critique", Working Papers, University of Fribourg (Faculty of Economic and Social Sciences), no. 272, June.


White, D. (1999) "Outlook Good, but Inflation a Worry", Financial Times, 10 September, in Deutsche Bundesbank (ed.) Auszüge aus Presseartikeln, no. 61, 15 September, p. 3.


