Essays in applied microeconometrics

Martin O'Connell

Thesis submitted to the Department of Economics
in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

University College London

London
July 2015

Declaration

No part of this thesis has been presented before to any university or college for submission as part of a higher degree. Chapter 2 was undertaken as joint work with Pierre Dubois and Rachel Griffith. Chapter 3 was undertaken as joint work with Rachel Griffith and Lars Nesheim. Chapter 4 was undertaken as joint work with Rachel Griffith and Helen Miller. Chapter 5 was undertaken as joint work with Rachel Griffith and Kate Smith.

Signature:

Date: July 23, 2015

Abstract

This thesis contains four self-contained papers in applied microeconometrics. Each uses an empirically tractable model of the behavior of economic agents, along with data on decisions taken by a sample of these agents, to learn about the determinants of choice and how they interact with the market environment.

The first chapter provides an introduction. Chapters 2 and 3 present models of consumer demand and firm competition in markets characterized by differentiated product oligopoly. In Chapter 2 we focus on junk food markets and study the effects on market equilibrium and welfare of banning junk food advertising. We model the impact of advertising on consumer demand in a flexible but empirically tractable way, we allow for the pricing response of firms in the market, and we carefully consider the impact of the ban on welfare.

Much of the literature using discrete choice models of consumer demand make strong assumptions about the nature of income effects. In Chapter 3 we discuss restrictions that these assumptions place on the shape of demand and on welfare effects, and we explore the empirical consequences of relaxing them in an application in which we simulate the introduction of an excise tax.

In Chapter 4 we consider how corporate income taxes effect where firms choose to legally own intellectual property. We estimate a discrete choice model of firm location choice and use the estimates to simulate the effects of the recent introduction of preferential tax treatment for income arising from patents.

The question we address in Chapter 5 is how well are consumers able to adjust their choice behavior to deteriorating economic conditions. We model consumers' grocery shopping behavior over the Great Recession and show that, despite big falls in expenditure, consumers succeeded in smoothing both their total calories and nutritional quality of those calories.

Acknowledgments

For the duration of my time as a PhD student, I have been employed as an economist at the *Institute for Fiscal Studies*. I am indebted to the many people involved at the institute who make it a thoroughly supportive, stimulating and productive environment in which to conduct research.

I am especially grateful to my supervisors - Richard Blundell and Rachel Griffith - who have offered continuous support, as well as intellectual inspiration. Rachel has also been my co-author throughout my time as a PhD student - a collaboration that I am in no doubt has been instrumental in developing my skills as a researcher. I am also very grateful to my other coathors - Pierre Dubois, Helen Miller, Lars Nesheim and Kate Smith - from whom I have learned a lot and with whom I have immensely enjoyed working.

The work presented in this thesis has been jointly funded by the Economic and Social Research Council through the Centre for the Microeconomic Analysis of Public Policy and Open Research Area grant 'Developing Robust Methods for Evaluating Policies in the Markets for Food and Nutrition' and by the European Research Council through the grant 'Microeconomic Analysis of Prices, Food and Nutrition'. I am grateful for this funding.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This thesis contains four self-contained papers in applied microeconometrics. Each addresses a question concerning how economic agents make choices, and in particular, how their choices respond to changes in the economic environment. In each paper economic theory provides a framework for how to think about decision making, but the ultimate answer to how choices are affected by the economic environment is not provided by theory alone. In each chapter we combine an empirically tractable model of decision making with rich micro data to address the question at hand.

1.1 Models of market equilibrium

Chapters 2 and 3 of this thesis sit squarely in the empirical industrial organization literature. In each case the question of interest concerns how a particular policy change will affect the equilibrium in a particular market and the welfare of the agents interacting in the market. Both chapters consider specific food markets in which the selection of products available to consumers are numerous, differentiated, and where consumers typically select one available option on each purchase occasion. These characteristics of the market mean that the most appropriate way to model consumer demand is by means of a discrete choice random utility model. Such models were originally introduced by McFadden (1974, 1981), and are based on an underlying utility maximization problem in which preferences are defined over characteristics of goods (Gorman (1980), Lancaster (1971)). One of the key advantages of this framework is that it provides the modeler with a way of estimating demand in markets where there are many products. The seminal contributions of Berry (1994) and Berry et al. (1995) showed how to allow for flexibility in price elasticities, by incorporating preference heterogeneity, and for unobserved characteristics in applications using market level date. In the work presented in this thesis we use micro data on the choices made by individual consumers to estimate the choice model.

In such markets the way in which the active, typically multi product, firms respond to policy change is likely to be an important determinant of the impact of the intervention. In particular, the firms are likely to be strategic players that will change their prices, and possibly other strategic variables, in response to policy change. Therefore it is important to allow for the possible response of firms. Following Nevo (2001) we couple demand estimates with an equilibrium pricing condition to estimate firms' marginal costs. This allows us to model how equilibrium prices will respond to changes in policy.

1.1.1 An advertising ban

The particular question addressed in Chapter 2 is: what are the effects of banning advertising in junk food markets? To answer this question, when specifying consumer demand, it is important to allow advertising to enter flexibly. In particular, if advertising of one brand is increased, demand for other brands in the market may rise or fall, and total demand for the product category may rise or fall. One contribution of this chapter is to introduce advertising into a discrete choice random utility model in a flexible but empirically tractable way, which *a priori* does not rule out any of these effects.

We also allow for the possibility that current advertising affects future demand. Therefore firms operating in the market play a dynamic game in which they set brand advertising budgets and product prices to maximize their discounted profits. In a realistic market setting in which many multi product firms play a dynamic game, computational problems can be prohibitive when solving for counterfactual market equilibria. However, we exploit the fact that we observe advertising state variables and we consider a counterfactual in which advertising is banned, which means in the counterfactual firms play a static pricing game. This enables us to fully account for the pricing response of firms to the ban. We find that, although banning advertising leads to a direct fall in demand, it also toughens price competition leading firms to lower their prices. The effect of lower equilibrium prices on demand is more than enough to offset the direct fall in demand from no advertising, meaning the overall effect of an advertising ban is to slightly increase the total quantity sold.

Ultimately, when considering such an intervention, we are interested in the associated welfare implications. If advertising enters directly into utility (Becker and Murphy (1993)) as a product characteristic, then welfare considerations are standard. An alternative view is that advertising is persuasive and acts to affect consumers' decisions but not their underlying utility; or in the words of Kahneman et al. (1997), advertising affects decision utility but not experience utility.

We show how to evaluate the consumer welfare effect of an advertising ban under this view of advertising; we find that the ban acts to raise consumer welfare, because in the counterfactual equilibrium consumers no longer make distorted choices and they benefit from facing lower prices.

We apply our model to the UK potato chip market. While our quantitative results depend on the specifics of the market, we argue that the substantive points we make apply more generally to junk food markets.

1.1.2 Income effects in random utility models

The vast majority of the existing literature on discrete choice random utility models make strong assumptions about the marginal utility of income. A common assumption is that the marginal utility of income is constant, formally ruling out income effects. In this case researchers often include income in the model as a "preference shifter" that linearly shifts the marginal impact of price on utility. An alternative commonly made assumption is that utility is linear in the log of money left for consumption outside the market. This incorporates income effects into the model, but it does so in a restrictive way; at the consumer level how income affects conditional utility is still dependant on just one parameter. This standard practice of making restrictive assumptions about the nature of income effects contrasts with the continuous choice demand literature which has concerned itself with allowing for increasingly general forms of income effects (Deaton and Muellbauer (1980), Banks et al. (1997), Lewbel and Pendakur (2009), Hausman and Newey (2014)).

In logit choice models constraining the marginal utility of income also imposes restrictions on the curvature of consumer level demands, which in turn places restrictions on the curvature of market demand. A series of theoretical papers (Seade (1985), Delipalla and Keen (1992) and Anderson et al. (2001), Weyl and Fabinger (2013)) show in stylized settings that demand curvature is a crucial determinant of pass-through of cost shocks and taxes to final consumer prices, and this is likely to be true in more complex realistic market settings.

The contribution of Chapter 3 is to explore the empirical importance of relaxing assumptions commonly placed on the marginal utility of income in logit demand models. We focus on the butter and margarine market, as this market comprises a small share of consumer expenditure and it is one in which *a priori* one may think income effects are not important, and we simulate the introduction of an excise tax. Under a number of different assumptions about the nature of income effects, we describe pass-through of the tax to consumer prices and the consumer welfare implications of the tax. Our principle finding is that failing to flexibly model how total consumer income or total expenditure affects demand is likely to lead to incorrect estimates of how demand and welfare effects vary across the income or total expenditure distribution.

1.2 Firm location choice

Intellectual property accounts for a growing share of firms' assets. It is more mobile than other forms of capital and could be used by firms to shift income offshore and to reduce their corporate income tax liability. In Chapter 4 we consider how corporate income taxes affect where firms choose to legally own intellectual property. We introduce the mixed or random coefficient logit model (widely exploited in the empirical industrial organization literature, including in the Chapter 2 and 3 of this thesis) to the public economics literature on firm location choice. To date papers in the firm location choice literature have tended to rely on binary choice and multinomial logit models. The approach taken in these papers has the drawback of placing unrealistic restrictions on substitution across locations in response to policy change. Our approach overcomes this drawback; we illustrate this by using the model to simulate the impact on location choice of recent reforms that have drastically cut the tax rates levied on patent income.

1.3 Grocery demand

A central question in economics is how do households insure themselves against shocks to their income and other shocks to their economic environment. There is a large literature that focuses on how households inter-temporally smooth their income through use of asset markets and family labor supply (including Hall and Mishkin (1982), Blundell et al. (2008), and Blundell et al. (2014)) to smooth their consumption. In Chapter 5 we focus on how households use intra-temporal substitution between money and time, and between different shopping basket characteristics, to mitigate the impact of shocks to their economic environment.

We focus on consumer grocery spending over the period of the Great Recession. We show that despite large increases in food prices and falls in real grocery expenditure, key aspects of consumption - total calories and their nutritional quality - remained remarkably smooth. To explain this we set out a model in which consumers can choose how much time to devote to grocery shopping in search of better deals (Stigler (1961)), as well as the size and characteristics of their grocery basket, and we use this to motivate our empirical investigation into how households adjusted their shopping behavior in response to toughening economic conditions.

We use longitudinal data to show that over the Great Recession households increased the intensity with which they purchase products on sale, indicating an increase in shopping effort. Aguiar and Hurst (2005, 2007) find a similar result in a different setting; they use cross-sectional variation in grocery purchases to argue that substitution from money to time is able to account for lower grocery expenditure among the retired. We also show that substitution between grocery basket characteristics was an important margin of adjustment. Households switched to cheaper calories, without compromising the nutritional quality of their shopping baskets, by switching towards more calories from generic products, and by switching to cheaper, but not necessarily less healthy, nutrients. Our overall conclusion is that consumers were adept at shifting their within period grocery consumption so as to mitigate the impact on the nutritional quality of their groceries of lower incomes and higher food prices.

There are a large number of exciting policy relevant research questions related to those addressed in this thesis. The modeling and data developed here should put us in a strong position to tackle a number of these questions. For example, in Chapter 2 we explore the effect on market equilibrium of an advertising ban. There is great interest in also understanding the effect on pricing and advertising of taxation. The framework we develop in Chapter 2 could potentially be used to study such a question. However, unlike when simulating an advertising ban, we would need to address the problem of conducting a counterfactual experiment in which firms play a dynamic game and consequently in which there are likely to be multiple equilibria. This is a challenging problem, but one for which, in slightly different settings, there has been recent advances in the literature (for instance Sweeting (2013)).

There is also much interest in how regulations impact on the equilibrium of oligopolistic markets. A recent example of regulation in food markets has been the adoption of market specific targets for the average nutrient content of sales. The motivation is to help reduce dietary intake of specific nutrients. However, in oligopolistic markets such an intervention could have implications for competition policy, because the nature of the target necessitates some degree of coordination among firm. The model of market equilibrium used in Chapters 2 and 3 can be extended to model such an intervention. In particular, like in these chapters, firms could be modeled as optimally choosing their strategic variables (in this case product prices and nutrient contents), but subject to an industry wide constraint on the set of permissible equilibrium nutrient values.

Chapter 2

The effects of banning advertising in junk food markets

2.1 Introduction

In this paper, we study the consequences of banning junk food advertising using a model of consumer demand and oligopoly supply in which multi product firms compete in prices and advertising budgets. We pay careful attention to the way that advertising affects demand, allowing advertising of one brand to potentially increase or decrease demand for other brands, and for past advertising to influence current demand. We use the model to simulate counterfactual market equilibria in which advertising is banned. At constant prices, banning advertising leads to a direct reduction in quantity demanded, but it also leads to tougher price competition and the decrease in equilibrium prices leads to an offsetting increase in quantity demanded.

We consider the possible impact of the ban on consumer health, by quantifying the impact it has on purchases of calories, saturated fat and salt. If one abstracts from the pricing response of the firms in the market, the ban has an unambiguously positive impact on health. Consumers respond both by switching out of the market, buying smaller pack sizes and by switching to healthier varieties that contain less saturated fat and less salt. However, the equilibrium pricing response of firms to the ban eliminates some of these health gains. Firms responds to the ban by lowering the prices of several of their products which causes demand to expand; the overall effect of the ban is actually to increase total calories, although, because consumers buy healthier varieties, purchases of saturated fat and salt do not increase.

We also consider the direct consumer welfare effect of the ban. The effect of banning advertising on consumer welfare depends on the view one takes on how advertising affects consumers' utility given its affect on consumers' decision making. Most advertising in junk food markets involves celebrity endorsements of established brands. If we view this as leading consumers to make choices that are inconsistent with their underlying preferences, with advertising not entering utility directly, the advertising leads consumers to maximize an objective function that differs from their true underlying utility, reflecting the distinction between decision utility and experienced utility (Kahneman et al. (1997)). Removing advertising leads to welfare gains because consumers' decisions are no longer distorted, and tougher price competition increases consumer welfare. On the other hand, if advertising enters utility directly then consumers are not made better off; the gain from lower prices is offset from the removal from the market of a valued product attribute.

Advertising is regulated in many markets (for example in the cigarette and tobacco and alcohol markets), with the aim of reducing consumption.[1] Attention has recently turned to using a similar policy tool to reduce the consumption of junk foods, particularly by children. The World Health Organization ((WHO, 2010)) published the recommendation that the "overall policy objective should be to reduce both the exposure of children to, and the power of, marketing of foods high in saturated fats, *trans*-fatty acids, free sugars, or salt". The medical literature has called for restrictions on advertising; for example, in a well cited paper, Gortmaker et al. (2011) state that "marketing of food and beverages is associated with increasing obesity rates", citing work by Goris et al. (2010), and say that it is especially effective among children, citing National Academies (2006) and Cairns et al. (2009).[2]

The aim of these interventions is to reduce consumption of junk foods. However, a ban on advertising could lead the market to expand or to contract. Brand advertising may be predatory, in which case its effect is to steal market share of rival products, or it might be cooperative, so that an increase in the advertising of one product increases demand for other products (Friedman (1983)). The impact on total market demand depends on the relative importance of these two effects.[3] In addition, firms are likely to respond to a ban on advertising by adjusting their prices, as the equilibrium prices with advertising are unlikely to be the same as in an equilibrium when advertising is not permitted.

[1]In other markets, such as pharmaceuticals and some professional services, the aim is more focused on consumer protection and information.

[2]In the UK, regulations ban the advertising of foods high in fat, salt or sugar during children's programming (see http://www.bbc.co.uk/news/health-17041347) and there have been recent calls to extend this ban (see http://www.guardian.co.uk/society/2012/sep/04/obesity-tv-junk-food-ads). In the US the Disney Channel has plans to ban junk food advertising (http://www.bbc.co.uk/news/world-us-canada-18336478).

[3]For example, Rojas and Peterson (2008) find that advertising increases aggregate demand for beer; while Anderson et al. (2012) show that comparative advertising of pharmaceuticals has strong business stealing effects and is detrimental to aggregate demand. Other papers show that regulating or banning advertising has led to more concentration, for example Eckard (1991), for cigarettes and Sass and Saurman (1995), for beer. Motta (2007) surveys numerous other studies.

To illustrate these effects we apply our model to the market for potato chips using novel data on purchases made both for consumption at home and purchases made on-the-go for immediate consumption by a sample of British consumers, combined with information on brand level advertising expenditure. The potato chips market is interesting because it is an important source of junk calories. However, we argue that it also representative of a broader set of junk food markets; like other junk food markets, it is dominated by a relatively small set of well established brands with advertising highly concentrated on the few most popular brands. Our results, therefore, are instructive about the effects of banning advertising in a broader range of junk food markets.

There is a large literature on how advertising affects consumer choice. Bagwell (2007) provides a comprehensive survey and makes a useful distinction between advertising as being persuasive, entering utility directly as a characteristics, or being informative. Much of the early literature on advertising focused on its persuasive nature (Marshall (1921), Braithwaite (1928), Robinson (1933), Kaldor (1950) and Dixit and Norman (1978)), where its purpose is to change consumer tastes. More recently, the behavioral economics and neuroeconomics literatures have explored the mechanisms by which advertising affects consumer decision making. Gabaix and Laibson (2006) consider models in which firms might try to shroud negative attributes of their products, while McClure et al. (2004) and Bernheim and Rangel (2004, 2005) consider the ways that advertising might affect the mental processes that consumers use when taking decisions (for example, causing a shift from the use of deliberative systems to the affective systems that respond more to emotional cues). This literature, in particular Dixit and Norman (1978), Bagwell (2007) and Bernheim and Rangel (2009), raises questions of how welfare should be evaluated, and particularly whether we should use preferences that are influenced by advertising or the "unadvertised self" preferences. Bernheim and Rangel (2009) argue that if persuasive advertising has no information content, choices based on the advertising cues are based on improperly processed information, and therefore welfare should be based on choice made under other conditions. In a theoretical paper, Glaeser and Ujhelyi (2010) adopts this perspective considering some advertising in the food market as misinformation that leads consumers to consume an unhealthy good excessively and argues that a quantity restriction on advertising can maximize welfare.

An alternative view of advertising is that it enters utility directly (see Becker and Murphy (1993) and Stigler and Becker (1977)). Consumers may like or dislike advertising, and advertising may act as a complement to other goods or characteristics that enters the utility function. The crucial feature that distinguishes this characteristic view of advertising from the persuasive view is how advertising affects consumer welfare. If advertising is viewed as a characteristic then it does not lead consumers to make decisions that are inconsistent with their true welfare, and consideration of the consumer welfare implications of banning advertising are analogous to those associated with removing or changing any other characteristic.

Another branch of the literature focuses on the role that advertising plays in providing information to consumers (as distinct from being persuasive). For instance, advertising may inform consumers about the quality or characteristics of a product (Stigler (1961) and Nelson (1995)), product price (for instance, see Milyo and Waldfogel (1999) who study the alcohol market), or about the existence and availability of products (see, inter alia, Sovinsky-Goeree (2008) on personal computers and Ackerberg (2001) and Ackerberg (2003) on distinguishing between advertising that is informative about product existence and prestige advertising in the yoghurt market). Although, as Anderson and Renault (2006) point out, firms may actually have an incentive to limit the informative content of adverts even when consumers are imperfectly informed (see also Spiegler (2006)). Studies of the tobacco bans of the 1970s show that these might have led to an increase in demand for cigarettes (see for example Qi (2013)). Bagwell (2007) provides a survey of the broader literature on advertising.

In junk food markets advertising consists of celebrity endorsements of established brands, and in our view they have little informative content. However, in considering the impact of an advertising ban on demand, prices and nutrients, we remain agnostic about the role of advertising, allowing it to shift demand in a flexible way. It is only when we consider the impact of a ban on consumer welfare that we need to be more specific about how advertising affects underlying consumer utility.

We characterize firms' observed behavior and show that we can identify marginal costs of all products without estimating the full dynamic game firms play when selecting their advertising budgets. While we could specify a fully dynamic oligopoly model that accounts for the potentially long lasting effects of advertising on firm behavior, we avoid many of the difficult computational problems that typical arise when using such models. In particular, in our setting firms compete in both prices and advertising; firms' strategies in prices and advertising are multidimensional and continuous with a very large set of state variables. If we required to estimate the dynamic parameters of the model, we would face a potentially intractable computational problem. However, we show that identifying marginal costs of products can be done using only static price optimality conditions along with observed values of the relevant state variables, without having to impose any restrictions on the likely multiplicity of advertising dynamic equilibrium strategies. Then, as we are interested in the counterfactual equilibrium in which advertising is banned, it is sufficient to focus on the static price first-order conditions in our simulations, since the dynamic effects of advertising on demand disappear in this counterfactual. This simplifies the problem, and allows us, in a realistic market setting, to consider the impact that an advertising ban will have on price competition.

The rest of the paper is structured as follows. In Section 2.2 we outline a model of consumer demand that is flexible in the ways that advertising enters, and allows for the possibility that advertising is cooperative so expands the market, or that it is predatory and so potentially contracts the market. Section 2.3 discusses firm competition in the market and outlines how we identify the unobserved marginal cost parameters of the model and how we simulate a counterfactual advertising ban. Section 2.4.1 describes the data used in our application to the UK potato chips market; a unique feature of our data is that we observe purchase decisions for consumption outside the home as well as at home. In this section we also describe the advertising in this market. Section 2.4.2 describes our estimates and Section 2.4.3 describes market equilibria with advertising and with an advertising ban implemented, emphasizing the effect the ban has on what nutrients consumers purchase. In Section 2.5 we discuss how to measure the consumer welfare effects of the ban and in Section 2.6 we show our main conclusions are robust to several modeling modifications. A final section summarizes and concludes.

2.2 Demand

We specify a demand model that is flexible in the way that advertising affects both individual and market level demand. We use a random utility discrete choice model in the vein of Berry et al. (1995), Nevo (2001) and Berry et al. (2004). We estimate the model on transaction level data. Berry and Haile (2010, 2014) show that identification of such multinomial choice models requires less restrictive assumptions with micro data, compared with when market level data alone are used.

2.2.1 Consumer choice model

Multi product firms offer brands ($b = 1, .., B$) in different pack sizes, indexed by s; a product index is defined by a (b, s) pair. Good $(0, 0)$ indexes the outside option of not buying potato chips. We index markets, which are defined as the period of time over which firms take pricing and advertising decisions, by t.

Let i index consumers. We observe individuals on two types of purchase occasion, food on-the-go and food at home, indexed by $\kappa \in \{1, 2\}$. On food on-the-go purchase occasions an individual buys a pack of potato chips for immediate consumption outside of the home; on food at home purchase occasions the main shopper in the household buys potato chips for future consumption at home.

Consumer i purchases the product that provides her with the highest payoff, trading off characteristics that increase her valuation of the product, such as tastiness, against characteristics that decrease her valuation, such as price and possibly 'unhealthiness'. Advertising could affect the weight the consumer places on different characteristics. This may occur because advertising itself is a characteristic that enters interactively with other product attributes, or advertising persuades consumers to place more or less weight on product characteristics than they would in the absence of advertising, or it could change the information the consumer has about the characteristics. For the moment we remain agnostic about whether advertising is playing a characteristic, persuasive or informative role; our main concern is identifying how advertising affects the consumer demand function. In Section 2.4.1 we discuss the nature of advertising in the UK potato chips market.

Products have observed and unobserved characteristics. A product's observed characteristics include its price (p_{bst}) and its nutrient characteristics (x_b). The nutrient characteristics might capture both tastiness, if consumers like the taste of salt and saturated fat, and the health consequences of consuming the product, which might reduce the payoff of selecting the product for some consumers. z_{bs} captures the product's pack size, and ξ_{ib} is an unobserved brand characteristic. A consumer's payoff from selecting a product also depends on an i.i.d. shock, ϵ_{ibst}. We also assume that there exists a set of advertising state variables, $\mathbf{a_t} = (\mathbf{a_{1t}}, ..., \mathbf{a_{Bt}})$, where $\mathbf{a_{bt}}$ denotes a brand b specific advertising vector of state variables, which may depend on the current brand advertising expenditures $\mathbf{e_t} = (e_{1t}, ..., e_{Bt})$ and possibly on all past advertising expenditures, since the start of the industry t_0, such that

$$\mathbf{a_t} = \mathcal{A}\left(\mathbf{e_t}, \mathbf{e_{t-1}}, .., \mathbf{e_{t_0}}\right).$$

As in Erdem et al. (2008), we assume that the dynamic effect of advertising on demand is such that the state advertising variables are equal to a geometric sum of current and past advertising expenditure:

$$\mathcal{A}\left(\mathbf{e_t}, \mathbf{e_{t-1}}, .., \mathbf{e_{t_0}}\right) = \sum_{n=0}^{t-t_0} \delta^n \mathbf{e_{t-n}},$$

which means that the dimension of the state space remains finite, since $\mathbf{a_t} = \mathcal{A}\left(\mathbf{a_{t-1}}, \mathbf{e_t}\right) = \delta \mathbf{a_{t-1}} + \mathbf{e_t}$. a_{bt} is akin to a stock of advertising goodwill that decays over time at rate δ, but that can be increased with expenditure e_{bt}. An alternative would be to specify that the state vector $\mathbf{a_t}$ depends on the vector of current brand advertising expenditures $\mathbf{e_t} = (e_{1t}, .., e_{Bt})$ and a maximum of L lagged advertising expenditures such that

$$\mathbf{a_t} = \mathcal{A}\left(\mathbf{e_t}, \mathbf{e_{t-1}}, .., \mathbf{e_{t-L}}\right).$$

In this case, the advertising state vector at the beginning of period t is not $\mathbf{a_{t-1}}$ but is $(\mathbf{e_{t-1}}, .., \mathbf{e_{t-L}})$. All of the following discussion can accommodate this case.

Letting $\bar{v}_{ibst} = \bar{v}_i(p_{bst}, \mathbf{a_t}, x_b, z_{bs}, \xi_{ib}, \epsilon_{ibst})$ denote the consumer's payoff from selecting product (b, s), the consumer will choose product (b, s) if:

$$\bar{v}_i(p_{bst}, \mathbf{a_t}, x_b, z_{bs}, \xi_{ib}, \epsilon_{ibst}) \geq \bar{v}_i(p_{b's't}, \mathbf{a_t}, x_{b'}, z_{b's'}, \xi_{ib'}, \epsilon_{ib's't}) \qquad \forall (b', s') \in \Omega_\kappa,$$

where Ω_κ denotes the set of products available on purchase occasion type κ. The i subscript on the payoff function indicates that we will allow coefficients to vary with observed and unobserved (through random coefficients) consumer characteristics. One of these characteristics is whether we observe the consumer making a purchase on-the-go for immediate consumption, or as part of a main shopping trip for future consumption at home. We therefore allow for the possibility that behavior and preferences might differ when a decision is made for immediate consumption compared to when it is made for delayed consumption. In our application this is important; for example, we find that consumers are more price sensitive when purchasing food on-the-go compared with when they purchase food to be consumed at home.

One of our aims in specifying the form of the payoff function is to allow changes in price and advertising to impact demand in a way that is not unduly constrained a priori. We therefore incorporate both observable and unobservable heterogeneity in consumer preferences. Many papers, including Berry et al. (1995), Nevo (2001) and Berry et al. (2004), have illustrated the importance of allowing for unobservable heterogeneity, in particular to allow flexible cross-price substitution patterns. While in differentiated markets it is typically reasonable to impose that goods are substitutes (lowering the price of one good increases demand for a second), it is not reasonable to impose that cross-advertising effects are of a particular sign. A priori we do not know whether more advertising of one brand increases or decreases demand for another brand. Therefore, we include advertising in consumers' payoff function in such a way that allows for the potential for both cooperative or predatory advertising.

In particular, let $d = \{1, ..., D\}$ index consumer groups that are based on the demographic group and purchase occasion type that a consumer belongs to. For instance, all consumers that are observed in the food at home sample and that have income below the median level, have occupations categorized as low skilled and have children will comprise one consumer group. In our application we have $D = 23$. We assume that consumer i's (from group d) payoff from selecting product (b, s) is given by:

$$\bar{v}_{ibst} = \alpha_{0i} p_{bst} + \psi_{0d} x_b + \eta_d z_{bs} + \xi_{ib} + \left[\lambda_i a_{bt} + \rho_i \left(\sum\nolimits_{l \neq b} a_{lt}\right) + \alpha_{1d} a_{bt} p_{bst} + \psi_{1d} a_{bt} x_b\right] + \epsilon_{ibst} \tag{2.1}$$

The first four terms capture the baseline impact of price, the nutrient characteristics, pack size and the unobserved brand effect on the payoff function. The coefficients on the nutrient characteristics and pack size are consumer group specific. The coefficient on price is consumer specific and given by:

$$\alpha_{0i} = \bar{\alpha}_{0d} + \nu_{di}^{\alpha} \tag{2.2}$$

$\bar{\alpha}_{0d}$ captures the baseline price coefficient (or, equivalently, negative marginal utility of income) for consumers that belong to group d. ν_{di}^{α} is a random variable that captures the deviation in the baseline price coefficient from the group d mean for consumer i. We allow for the distribution of the random coefficient ν_{di}^{α} to vary across consumer groups, meaning we include D random coefficients on price. We also write the unobserved brand effect with an i subscript because we will include a random coefficient on a subset of these.

The portion of equation (2.1) in square brackets captures the impact of advertising on the payoff function. Advertising enters the payoff function in three distinct ways. We allow advertising to enter directly in levels. This can also be viewed as allowing advertising to shift the weight the consumer places on the brand characteristic; the coefficients λ_i and ρ_i can be interpreted as capturing the extent to which time variation in the own brand and competitor advertising state vectors shift the weight consumers place on the brand. We allow advertising to interact with price; the coefficient α_{1d} allows the marginal effect of price on the payoff function to shift with advertising (as in Erdem et al. (2008)). We also allow advertising to interact with the nutrient characteristics; the coefficient ψ_{1d} allows the marginal effect of the nutrient characteristics on the payoff function to shift with advertising.

We allow all of the coefficients on advertising to vary across consumers. This rich variation is important, because it allows for the possibility that different consumers have different levels of exposure to the advertising states. It also allows for the possibility of differential demand responsive to advertising (for a given level exposure). In particular, we specify the coefficients on the own brand and competitor advertising effects as:

$$\lambda_i = \bar{\lambda}_d + \nu_{di}^{\lambda} \tag{2.3}$$

$$\rho_i = \bar{\rho}_d + \nu_{di}^{\rho}, \tag{2.4}$$

and we allow the interactive advertising effects to vary across the D consumer groups. In each case the baseline impact is consumer group specific. So, for instance, $\bar{\lambda}_d$ captures the direct level impact of more brand advertising on the payoff functions for consumer belonging to group d. As there may still be some residual heterogeneity in the effect of advertising on the payoff function within consumer groups, due to differences in exposure and sensitivities to advertising, we also allow for random coefficients on the level effect of own brand and competitor advertising, ν_{di}^{λ} and ν_{di}^{ρ}. As with the price random coefficient; we allow the distributions of ν_{di}^{λ} and ν_{di}^{ρ} to vary across consumer groups.

Our approach to allowing for differential advertising exposures (and responses) is to allow for a very high degree of observable and unobservable heterogeneity; we include $2D$ (i.e 46) random coefficient on advertising and $2D$ non-random coefficients on advertising interacted with product attributes. We believe that this rich heterogeneity will capture variation in advertising exposure across consumers. In the robustness section (Section 2.6) we show that allowing a measure of consumer TV exposure (the vast majority of junk food advertising is done on TV) to shift all the advertising coefficients yields very similar results.

Without further restrictions, we cannot separately identify the baseline nutrient characteristics coefficient (ψ_{0d}) from the unobserved brand effect (ξ_{ib}), although we can identify the sum of the two effects. To simulate the equilibrium and welfare effects of an advertising ban, it is not necessary to separately estimate these coefficients. However, we are able to recover the baseline effect of the nutrient characteristics on the payoff function under the additional assumption that the nutrient characteristics are mean independent from the unobserved brand characteristic, using an auxiliary minimum distance estimation involving the set of estimated brand effects.

We allow for the possibility that the consumer chooses not to purchase potato chips; the payoff from selecting the outside option takes the form:

$$\bar{v}_{i00t} = \zeta_{d0t} + \epsilon_{i00t}.$$

We allow the mean utility of the outside option for each consumer group, ζ_{d0t} to change over time. In particular, we allow it to change from year to year and seasonally.

Assuming ϵ_{ibst} is i.i.d. and drawn from a type I extreme value distribution, and denoting $\boldsymbol{\zeta_t}$ the vector of time specific state variables affecting demand (time varying shocks to the outside good utility), the probability that consumer i buys product (b, s) in period (market) t is:

$$s_{ibs}(\mathbf{a_t}, \mathbf{p_t}, \boldsymbol{\zeta_t}) = \tag{2.5}$$
$$\frac{\exp\left(\alpha_{0i}p_{bst} + \psi_{0d}x_b + \eta_d z_{bs} + \xi_{ib} + \left[\lambda_i a_{bt} + \rho_i(\sum_{l\neq b} a_{lt}) + \alpha_{1d}a_{bt}p_{bst} + \psi_{1d}a_{bt}x_b\right]\right)}{\exp(\zeta_{d0t}) + \sum_{(b',s')\in\Omega_\kappa} \exp\left(\alpha_{0i}p_{b's't} + \psi_{0d}x_{b'} + \eta_d z_{b's'} + \xi_{ib'} + \left[\lambda_i a_{b't} + \rho_i(\sum_{l\neq b'} a_{lt}) + \alpha_{1d}a_{b't}p_{b's't} + \psi_{1d}a_{b't}x_{b'}\right]\right)}$$

2.2.2 Price and advertising effects

The specification of the payoff function described in the preceding section responds to both the need for flexibility and the need for parsimony. We allow for the possibility that an increase in the advertising state variable for one brand, b, increases demand for another brand b', in which case the advertising is cooperative with respect to brand b'. We also allow for the alternative possibility that it decreases demand for brand b', in which case the advertising is predatory with respect to brand b'. The size of the total market can expand or contract in response to an increase in the brand b advertising state. It is important that we include advertising in the model flexibly enough to allow for the possibility of these different effects.

With our specification of the consumer choice model the marginal impact of a change in an advertising state variable of one brand ($b > 0$) on the individual level choice probabilities is given by:

$$\begin{aligned}
\frac{\partial s_{ibst}}{\partial a_{bt}} =& s_{ibst}\left[\tilde{\lambda}_{ibst} - \rho_i(1 - s_{i00t}) - \sum\nolimits_{s'\in K_b}(\tilde{\lambda}_{ibs't} - \rho_i)s_{ibs't}\right] \\
\frac{\partial s_{ib'st}}{\partial a_{bt}} =& s_{ib'st}\left[\rho_i s_{i00t} - \sum\nolimits_{s'\in K_b}(\tilde{\lambda}_{ibs't} - \rho_i)s_{ibs't}\right] \qquad \text{for } b' \neq (0, b) \\
\frac{\partial s_{i00t}}{\partial a_{bt}} =& - s_{i00t}\left[\rho_i(1 - s_{i00t}) + \sum\nolimits_{s'\in K_b}(\tilde{\lambda}_{ibs't} - \rho_i)s_{ibs't}\right],
\end{aligned}$$

where $\tilde{\lambda}_{ibst} = \lambda_i + \alpha_{1d}p_{bst} + \psi_{1d}x_b$ and K_b denotes the set of all pack sizes s that brand b is available in.

The interaction of the advertising state variable a_{bt} with price and the nutrient characteristics, and the possibility that competitor advertising directly enters the payoff function are important in allowing for a flexible specification. If we do not allow competitor advertising to affect demand (imposing $\rho_i = 0$), and do not allow advertising to affect the consumer's responsiveness to price or nutrients (imposing $\alpha_{1i} = 0$ and $\psi_{1i} = 0$), then we directly rule out cooperative advertising and market contraction. In this case, the marginal impacts would be, for $b > 0$:

$$\begin{aligned}
\frac{\partial s_{ibst}}{\partial a_{bt}} =& \lambda_i s_{ibst} \left[1 - \sum\nolimits_{s' \in K_b} s_{ibs't}\right] \\
\frac{\partial s_{ib'st}}{\partial a_{bt}} =& -\lambda_i s_{ib'st} \left[\sum\nolimits_{s' \in K_b} s_{ibs't}\right] \qquad \text{for } b' \neq b.
\end{aligned}$$

In this restricted model, in order for advertising to have a positive own effect (so $\partial s_{ibst}/\partial a_{bt} > 0$) we require $\lambda_i > 0$. In this case, advertising at the consumer level is predatory (since $\partial s_{ib'st}/\partial a_{bt} < 0$), and it necessarily leads to market expansion (since $\partial s_{i00t}/\partial a_{bt} < 0$).

Allowing α_{1d} and ψ_{1d} to be non-zero, but with no competitor advertising effect ($\rho_i = 0$), makes the model more flexible. However, it will in general also restrict, at the consumer level, advertising to be predatory, and to lead to market expansion if own advertising increases own market share ($\partial s_{ibst}/\partial a_{bt} > 0$). However by allowing $\rho_i \neq 0$ we can capture more general effects. This does come at the expense of making direct interpretation of the advertising coefficients more difficult, for example, we can have $\lambda_i < 0$ but nonetheless have advertising have a positive own demand effect. However, it is straightforward to use estimates of the demand model to shut off advertising of one brand, to determine the effect it has on demands.

The interaction of advertising with price also allows advertising to have a direct impact on consumer level price elasticities. In particular, the our specification yields consumer level price elasticities given by, for $b > 0$ and $s > 0$:

$$\begin{aligned}
\frac{\partial \ln s_{ibst}}{\partial \ln p_{bst}} &= (\alpha_{0i} + \alpha_{1d} a_{bt})\,(1 - s_{ibst}) p_{bst} \\
\frac{\partial \ln s_{ib's't}}{\partial \ln p_{bst}} &= -\,(\alpha_{0i} + \alpha_{1d} a_{bt})\, s_{ibst} p_{bst} \qquad \text{for } b' \neq b \text{ or } s' \neq s.
\end{aligned}$$

Hence, advertising impacts consumer level price elasticities in a flexible way, through its impact on choice probabilities and through its impact on the marginal effect of price on the payoff function captured by α_{1d}.

2.2.3 Market demand

Market level demand is obtained by aggregating individual demands. The inclusion of rich observed and unobserved consumer heterogeneity means that flexibility in individual demand will translate into even more flexibility in market level demand. We consider firms to take pricing and advertising decisions each month t. We measure the potential size of the potato chips market (or maximum number of potato chips that could be purchased) as being equal to the number of shopping occasions on which snacks were purchased, denote this M_t. This definition of the market size implies that we assume that changes in pricing or advertising in the potato chips market may change consumers' propensity to buy potato chips, but not their propensity to go shopping to buy a snack product. We model the share of the potential market accounted for by purchases of product (b, s), by averaging over the individual purchase probabilities given by equation (2.5).

To aggregate individual choice probabilities into market shares we assume that random coefficients $\pi_i^u = (\alpha_{0i}, \lambda_i, \rho_i, \xi_i)$ are i.i.d. across consumers, within each consumer type group. As seen in Section 2.2.1, we allow the mean and variance of the random coefficient to vary with consumer type group, d; this captures both consumer characteristics and whether the purchase occasion is for food in or food on-the-go. We integrate over consumers' observed and unobserved preferences; under this assumption the share of the potential market accounted for by product (b, s) is given by:

$$s_{bs}(\mathbf{a_t}, \mathbf{p_t}, \boldsymbol{\zeta_t}) = \int s_{ibs}(\mathbf{a_t}, \mathbf{p_t}, \boldsymbol{\zeta_t}) dF(\pi_i^u, \pi_i^o, d). \tag{2.6}$$

where $F(\pi_i^u, \pi_i^o, d)$ is the joint cumulative distribution function of random coefficients π_i^u, fixed coefficients $\pi_i^o = (\psi_{0d}, \alpha_{1d}, \psi_{1d})$ and demographic groups d.

The assumption that π_i^u are i.i.d. across consumers, within each consumer type group, guarantees that the market share function $s_{bs}(., ., .)$ is not time dependent. A generalization, where the distribution of observed preference shifters d changes over time, is straightforward and would simply mean that the parameters of this distribution at time t would be an additional argument of the demand state variables $\boldsymbol{\zeta_t}$ in $s_{bs}(., ., .)$.

2.2.4 Identification

A common concern in empirical demand analysis is whether the ceteris paribus impact of price on demand is identified. In the industrial organization literature the most common concern is that price is correlated with an unobserved product effect (either some innate unobserved characteristic of the product or some market specific shock to demand for the product); failure to control for the unobserved product effect will mean that we cannot identify the true effect of price on demand. We use the identification strategy suggested by Bajari and Benkard (2005), exploiting the richness of our micro purchase and our advertising data, and the fact that the UK retail food market is characterized by close to national pricing.[4]

The use of individual transaction level data, coupled with the lack of geographic variation in pricing, means in our context that concerns over the endogeneity of price translates into whether differences in national price, either across products or through time, are correlated with the *individual level* errors (ϵ_{ibst}), conditional on all other characteristics included in the model.

[4] In the UK most supermarkets implement a national pricing policy following the Competition Commission's investigation into supermarket behavior (Competition Commission (2000)).

We observe and control for all relevant brand advertising in the market, alleviating concerns that omitted marketing activities are correlated with prices. We also control for time effects and seasonality of aggregate potato chip demand. Hence, we are able to exploit time series variation in product level prices, conditional on brand advertising and time effects in aggregate potato chip demand. A second common endogeneity concern is that some unobserved characteristic of products is not adequately captured by the model and firms, which set prices based in part on the demand they face, will set prices that are correlated with the unobserved characteristic. A strength of our data is that we observe barcode level transactions, and we are therefore able to model demand for products that are defined more finely than brands (in particular each brand is available in a variety of pack sizes). We control for both brand effects, which capture unobserved characteristics of brand, as well as pack size. So a second source of price variation we exploit is differences across brand in how unit price varies across pack size (non-linear pricing). Taken together we believe that national pricing and marketing decisions, consumer level demand, and the inclusion of aggregate time effects, advertising and brand effects deal with the typical sources of endogeneity of prices.

A related issue is whether we are able to identify the ceteris paribus effect of advertising on demand. Like pricing decisions, in the UK, advertising decisions are predominantly taken at the national level; the majority of advertising is done on national television, meaning that all consumers are potentially subject to the same vector of advertising. Of course, individual exposures to advertising may vary, and we allow for this by including very rich observable and unobservable heterogeneity in how advertising affects demand. Given this, we think it is reasonable to assume that the individual errors in our demand model are not correlated with national brand advertising.

Although we feel that the combination of our rich data and specific features of the UK market make our assumption that price and advertising are uncorrelated with demand shocks a reasonable one, some doubt may remain about the assumption's validity. In the robustness section (Section 2.6) we therefore present results obtained from implementing a control function approach. We instrument for product prices using prices from a different market (see Hausman (1996)) and instrument for contemporaneous advertising using advertising from a different product category (which will vary with the price of advertising but not with demand shocks in the potato chip market). The results are qualitatively similar.

In the counterfactual, we use our model to predict demand in the absence of advertising. An associated concern related to this exercise is that we may be predicting demand outside the range of variation in our data and that this may give a misleading picture of the true effects of an advertising ban. However, a number of brands in the market, in particular the generic supermarket brands, effectively do not advertise so for these brands we observe the advertising state variable at (very close to) zero. Other brands sometimes have long periods without advertising, meaning that if sufficiently old past histories of advertising do not affect the state variables a_{bt}, for these brands also we are close to observing states that are similar to what they would be if advertising was banned.

2.3 Supply model and counterfactual advertising ban equilibrium

If firms are forward looking, they will account for the fact that advertising decisions taken in one period affect demand contemporaneously and in the future. In addition, these decisions will affect current and future demand of other firms in the market. Therefore, when setting their price and advertising budgets, firms will play a dynamic oligopoly game. In any equilibria to this game profit maximizing firms will form dynamic strategies that may be very complex. The applied literature on dynamic games has typically dealt with such complicated dynamic games by considering Markov Perfect Equilibrium and by focusing on relatively stylized settings (see, for instance, Maskin and Tirole (1988) and Ericson and Pakes (1995)). In Appendix Section A.1 we outline how such modeling can be applied to our market setting in which multi products make dynamic advertising decisions.

For our purposes though, it is not necessary to specify fully the dynamic oligopoly game. We can use the fact that, in our demand model, product prices are an argument of current demand and profits, but not future demand and profits. In addition, prices do not directly influence the evolution of the advertising state variables. Therefore, conditional on the state variables, prices are determined in a static equilibrium in which firms choose prices to maximize current static profits. Then, given we observe the advertising states (which are simply functions of current and past advertising expenditures), we can use the static price conditions to identify firms' marginal costs.

In particular, let firms in the market be indexed by $j = 1, ..., J$, and denote the set of products (brand-pack size combinations (b, s)) produced by firm j, F_j, and the set of brands owned by firm j, B_j. Conditional on the advertising state variables $\mathbf{a_t}$ firm j, at time t, chooses the prices of its products to maximize flow variables profits:

$$\sum\nolimits_{(b,s)\in F_j} (p_{bst} - c_{bst})\, s_{bs}\left(\mathbf{a_t}, \mathbf{p_t}, \boldsymbol{\zeta_t}\right) M_t - \sum\nolimits_{b\in B_j} e_{bt}, \tag{2.7}$$

where c_{bst} is the marginal cost of product (b,s) at time t and $\sum_{b\in B_j} e_{bt}$ is the total advertising expenditure by firm j during period t. The set of price first order conditions for firm j are then:

$$s_{bs}\left(\mathbf{a_t},\mathbf{p_t},\boldsymbol{\zeta_t}\right)+\sum_{(b',s')\in F_j}\left(p_{b's't}-c_{b's't}\right)\frac{\partial s_{b's'}\left(\mathbf{a_t},\mathbf{p_t},\boldsymbol{\zeta_t}\right)}{\partial p_{bst}}=0 \qquad \forall (b,s)\in F_j. \tag{2.8}$$

With knowledge of the shape of demand, and observations on the advertising states and prices, we can use the set of price first order conditions (2.8) for all firms to identify marginal costs, provided the system of equations is invertible, which will be the case if goods are "connected substitutes" as in Berry and Haile (2014).

A second set of conditions characterizing the optimal choice of advertising flows as a function of past state variables may exist (see Appendix Section A.1). However, we do not need to appeal to these conditions to identify marginal costs; the price first order conditions are sufficient for this purpose.

We need to infer the vector of marginal costs in each market in order to conduct the counterfactual experiment of a ban on advertising. It is straightforward to show that, following the introduction of an advertising ban, equilibria will satisfy the per period Nash-Bertrand conditions of profit maximization, whatever the beliefs of firms about whether the regulatory change is permanent or not. We assume that technical conditions on the demand shape are satisfied to guarantee uniqueness of a Nash equilibrium. In the absence of advertising, the new price equilibrium $\mathbf{p_t^0}$ must be such that, for all (b,s) and j,

$$s_{bs}\left(\mathbf{0},\mathbf{p_t^0},\boldsymbol{\zeta_t}\right)+\sum_{(b',s')\in F_j}\left(p^0_{b's't}-c_{b's't}\right)\frac{\partial s_{b's'}\left(\mathbf{0},\mathbf{p_t^0},\boldsymbol{\zeta_t}\right)}{\partial p_{bst}}=0, \tag{2.9}$$

where

$$s_{bs}(\mathbf{0},\mathbf{p_t^0},\boldsymbol{\zeta_t})=\int s_{ibs}(\mathbf{0},\mathbf{p_t^0},\boldsymbol{\zeta_t})dF(\pi_i^u,\pi_i^o,d) \tag{2.10}$$

is the market level demand for product (b,s) when advertising stocks are all zero and at prices $\mathbf{p_t^0}$. We could easily obtain the counterfactual equilibrium for any other exogenously fixed levels of advertising state variables. For example, we could consider a period t_0 level of advertising state variables and simulate the subsequent period equilibria, where the advertising state vector decreases over time due to the decay of advertising, since $a_{t_0+L}=\delta^L a_{t_0}$ is decreasing in L and converging to zero.

To evaluate the impact of an advertising ban we solve for the counterfactual pricing equilibrium, defined by the equations (2.9) and (2.10), in each market and compare the quantities, prices and profits relative to the equilibrium played prior to the ban (the outcome of which we observe).

The price equilibrium under an advertising ban will be different from the observed one because of the change in the demand shape. In particular, advertising state variables affect the price first order conditions in two ways. They affect the demanded quantities through the way $s_{bs}\left(\mathbf{a_t}, \mathbf{p_t}, \boldsymbol{\zeta_t}\right)$ depends on $\mathbf{a_t}$ and they affect the price derivatives of market shares through the way $\frac{\partial s_{b's'}(\mathbf{a}, \mathbf{p_t}, \boldsymbol{\zeta_t})}{\partial p_{bst}}$ depends on $\mathbf{a_t}$. In Section 2.2.2 we highlighted that our demand model allows advertising to have flexible effects on consumer demand levels and slopes. The inclusion of rich consumer heterogeneity in the model translates into an even more flexible relationship between advertising and the shape of market demand.

In the case of a single product firm, using the implicit function theorem, it is straightforward to express the partial derivative of the equilibrium price $p_{bst}(\mathbf{a_t}, \boldsymbol{\zeta_t})$ as a function of advertising of any brand:

$$\frac{\partial p_{bst}(\mathbf{a_t}, \boldsymbol{\zeta_t})}{\partial a_{b't}} = \frac{\frac{\partial s_{bs}(\mathbf{a_t}, \mathbf{p_t}, \boldsymbol{\zeta_t})}{\partial a_{b't}} + \left(p_{bst}(\mathbf{a_t}, \boldsymbol{\zeta_t}) - c_{bst}\right) \frac{\partial^2 s_{bs}(\mathbf{a_t}, \mathbf{p_t}, \boldsymbol{\zeta_t})}{\partial p_{bst} \partial a_{b't}}}{-\left[2\frac{\partial s_{bs}(\mathbf{a_t}, \mathbf{p_t}, \boldsymbol{\zeta_t})}{\partial p_{bst}} + \left(p_{bst}(\mathbf{a_t}, \boldsymbol{\zeta_t}) - c_{bst}\right) \frac{\partial^2 s_{bs}(\mathbf{a_t}, \mathbf{p_t}, \boldsymbol{\zeta_t})}{\partial p_{bst}^2}\right]}.$$

In general the denominator and the price-cost margin will be positive. Therefore the sign of the change in price with respect to one brand advertising state variable will depend on the relative change of the market share and its price slope with respect to the brand advertising. Consider the case of $b' = b$. If the own demand effect of advertising is positive ($\frac{\partial s_{bs}(\mathbf{a_t}, \mathbf{p_t}, \boldsymbol{\zeta_t})}{\partial a_{bt}} > 0$), then a marginal increase in advertising will raise the equilibrium price as long as the effect of advertising is not to increase the price slope of demand by too much (i.e. as long as $\frac{\partial^2 s_{bs}(\mathbf{a_t}, \mathbf{p_t}, \boldsymbol{\zeta_t})}{\partial p_{bst} \partial a_{bt}}$ is not too negative). However, the effect of an advertising ban on equilibrium prices will be more complicated as it will depend on the effect of (non-marginal) changes in all advertising state variables. The effect of a ban on the price of product (b, s) can written as:

$$p_{bst}\left(\mathbf{0}, \boldsymbol{\zeta_t}\right) - p_{bst}\left(\mathbf{a_t}, \boldsymbol{\zeta_t}\right) = -\sum_{b'=1}^{B} \int_0^{a_{b't}} \frac{\partial p_{bs}\left(0, ., 0, x, a_{b'+1t} .., a_{Bt}, \boldsymbol{\zeta_t}\right)}{\partial x} dx.$$

2.4 Application

2.4.1 Data

We apply our model to the UK market for potato chips. This is an important source of junk food calories. In the US the potato chips market was worth \$9 billion in 2013, and 86% of people consumed some potato chips. The UK potato chips market had an annual revenue of more than £1.2 billion in 2010 with 84% of consumers buying some potato chips.[5]

[5]For the size of the US market see http://www.marketresearch.com/MarketLine-v3883/Potato-Chips-United-States-7823721/ ; the size of the UK market see http://www.marketingmagazine.co.uk/article/1125674/sector-

This market shares several important characteristics with other junk food markets, which make our results of broader interest. In particular, there are a small number of large firms that sell multiple well established brands and that have large advertising budgets, while advertising is mainly in the form of celebrity endorsement or other types that contain little factual information on the characteristics of the product.

We use two sources of data - transaction level purchase data from the market research firm Kantar, and advertising data from AC Nielsen.

Purchase data

The purchase data are from the Kantar World Panel for the period June 2009 to October 2010. Our data are unusual in that we have information on households' purchases for food at home *and* individuals' purchases for food on-the-go. For each household we observe *all* food purchases made and brought into the home (we refer to these as "food at home" purchases). We also use a sample of individuals drawn from these households that record all food purchases made for consumption "on-the-go" (we refer to these as "food on-the-go" purchases) during the same period. Food at home purchases are by definition made for future consumption (the product has to be taken back home to be recorded), while food on-the-go purchases are made for immediate consumption. Individuals participating in the on-the-go panel include both adults and children aged 13 or older.

We use information on 260,682 transactions over the period June 2009 to October 2010; this includes 161,392 food at home purchase occasions and 99,290 food on-the-go purchase occasions, made by 2,870 households and 2,289 individuals. We define a purchase occasion as a week. For the food at home segment this is any week in which the household records buying groceries; when a household does not record purchasing any potato chips for home consumption we say it selected the outside option in this segment. Potato chips are purchased on 41% of food at home purchase occasions. For the food on-the-go segment a purchase occasion is any week in which the individual records purchasing any food on-the-go; when an individual bought food on-the-go, but did not purchase any potato chips, we say they selected the outside option. Potato chips are purchased on 27% of food on-the-go purchase occasions. From other data we know that 14% of potato chips are bought on-the-go, with the remaining share purchased for food at home (Living Cost and Food Survey).

insight-crisps-salty-snacks ; and for the number of people who consume potato chips in each country see http://us.kantar.com/business/health/potato-chip-consumption-in-the-us-and-globally-2012/.

We define a potato chip product as a brand-pack size combination.[6] Potato chips for consumption at home are almost entirely purchased in large supermarkets as part of the households' main weekly shop, whereas those for consumption on-the-go are almost entirely purchased in small convenience stores. The set of products available in large supermarkets (for food at home) differs from the set of products available in convenience stores (for food on-the-go). Some brands are not available in convenience stores (for example, generic supermarket brands), and purchases made at large supermarkets are almost entirely large or multi-pack sizes, while food on-the-go purchases are almost always purchases of single packs. We restrict the choice sets in each segment to reflect this. This means that the choice sets for food at home and on-the-go occasions do not overlap; most brands are present in both segments, but not in the same pack size. Table 2.1 shows the set of products available and the market shares in each market segment. The table makes clear that Walkers is, by some distance, the largest firm in the market - its products account for 46% of all potato chips sold in the food at home segment and 55% of that sold in the food on-the-go segment. For each product we compute the transaction weighted mean price in each of the 17 months (or markets). Table 2.1 shows the mean of these market prices. We use these market prices in demand estimation.

[6]Potato chips are available in a variety of different flavors, for example, salt and vinegar or cheese and onion are popular flavors. We have this information, but we do not distinguish between these products because neither price nor advertising varies within product across flavors. For our purposes it is the choice that a consumers makes between brand and pack size that is relevant.

Table 2.1: *Quantity share and mean price*

Segment:	**Food at home**		**Food on-the-go**	
	Quantity Share	Price (£)	Quantity Share	Price (£)
Walkers	45.66%		54.70%	
Walkers Regular:34.5g			27.40%	0.45
Walkers Regular:50g			7.21%	0.64
Walkers Regular:150-300g	1.77%	1.26		
Walkers Regular:300g+	23.97%	2.79		
Walkers Sensations:40g			2.05%	0.62
Walkers Sensations:150-300g	0.43%	1.26		
Walkers Sensations:300g+	1.80%	2.58		
Walkers Doritos:40g			4.74%	0.54
Walkers Doritos:150-300g	1.30%	1.21		
Walkers Doritos:300g+	3.29%	2.50		
Walkers Other:<30g			4.37%	0.45
Walkers Other:30g+			8.94%	0.61
Walkers Other:<150g	0.69%	1.24		
Walkers Other:150-300g	3.73%	1.77		
Walkers Other:300g+	8.67%	3.17		
Pringles	6.86%			
Pringles:150-300g	1.33%	1.11		
Pringles:300g+	5.53%	2.61		
KP	19.51%		23.21%	
KP:50g			23.21%	0.52
KP:<150g	0.21%	0.86		
KP:150-300g	4.81%	1.19		
KP:300g+	14.49%	2.39		
Golden Wonder	1.50%		4.24%	
Golden Wonder:<40g			3.14%	0.39
Golden Wonder:40-100g			1.10%	0.73
Golden Wonder:<150g	0.10%	1.27		
Golden Wonder:150-300g	0.25%	1.41		
Golden Wonder:300g+	1.15%	2.78		
Asda	3.34%			
Asda:<150g	0.09%	0.94		
Asda:150-300g	0.90%	0.95		
Asda:300g+	2.36%	2.29		
Tesco	6.47%			
Tesco:<150g	0.18%	0.82		
Tesco:150-300g	1.78%	0.91		
Tesco:300g+	4.51%	2.07		
Other	16.65%		17.84%	
Other:<40g			12.03%	0.50
Other:40-100g			5.81%	0.66
Other:<150g	0.94%	1.05		
Other:150-300g	3.94%	1.31		
Other:300g+	11.77%	2.57		

Notes: Quantity share refers to the quantity share of potato chips in the segment accounted for by that product. Price refer to the mean price across markets.

We are particularly interested in the nutrient characteristics of the products. Table 2.2 shows the main nutrients in potato chips. In our baseline demand estimation we control for the nutrient characteristics using an index that combines the individual nutrients into a single score and that is used by UK government agencies. It is based on the nutrient profile model developed by Rayner et al. (2005) (see also Rayner et al. (2009) and Arambepola et al. (2008))

and is used by the UK Food Standard Agency, and by the UK advertising regulator Ofcom to determine the healthiness of a product. For potato chips the relevant nutrients are the amount of energy, saturated fat, sodium and fiber that a product contains per 100g. Products get points based on the amount of each nutrient they contain; 1 point is given for each 335kJ per 100g, for each 1g of saturated fat per 100g, and for each 90mg of sodium per 100g (or, equivalently, 0.225g of salt per 100g). Each gram of fiber per 100g reduces the score by 1 point. The UK Food Standard Agency uses a threshold of 4 points or more to define 'less healthy' products, and Ofcom has indicated this is the relevant threshold for advertising restrictions (Ofcom (2007)). In the robustness section we show that our main results are robust to including individual nutrients in our demand model in place of the single index measure.

Table 2.2 also shows the nutrient profile. There is considerable variation across brands; Walkers Regular has the lowest score (10), and the brand Pringles and KP have the highest score (18). This is a large difference. To give some context, if all other nutrients were the same then an 8g difference in saturated fat (per 100g or product) would lead to a difference of 8 points in the nutrient profile score; in the UK the guideline daily amount of saturated fat is 20g per day for woman and 30g per day for men. Note also that potato chips lie far above the 'less healthy' threshold of 4 and that reformulation to bring them below the threshold is unlikely.

Table 2.2: *Nutrient characteristics of brands*

Brand	Nutrient profiling score	Energy (kj per 100g)	Saturated fat (g per 100g)	Salt (g per 100g)	Fiber (g per 100g)
Walkers Regular	10	2168	2.56	1.48	4.04
Walkers Sensations	11	2021	2.16	1.78	4.25
Walkers Doritos	12	2095	2.86	1.65	3.02
Walkers Other	15	2017	2.50	2.04	3.14
Pringles	18	2160	8.35	1.55	2.74
KP	18	2159	5.87	2.10	2.70
Golden Wonder	16	2127	4.03	2.30	3.77
Asda	15	2125	4.13	1.88	3.31
Tesco	15	2142	4.63	1.92	3.57
Other	12	2081	3.84	1.75	4.06

Notes: See text for definition of the nutrient profiling score.

Table 2.3 provides details of the numbers of households of each type, the number of individuals making food on-the-go decisions and the number of purchase occasions. Households and individuals can switch between demographic groups, for example if a child is born in a household, or if a grown up child turns 18.

We allow all coefficients, including the distribution of the random coefficients, to vary across the demographic groups shown in Table 2.3. Households are distinguished along three characteristics: (i) household composition, (ii) skill or education level of the head of household, based on socio-economic status, and (iii) income per household member.

Table 2.3: *Household types*

Demographic group			Number of		Number of purchase occasions	
			households	individuals	food at home	food on-the-go
Composition	skill level	income				
No children	high	high	472	344	22718	14388
		medium	308	236	13149	8364
		low	289	251	13330	8213
	low	medium-high	215	164	10185	6684
		low	343	255	16141	8530
Pensioners			270	144	14345	6003
Children	high	high	409	342	20435	12776
		medium	315	258	14294	8409
		low	165	136	7087	4490
	low	medium-high	322	268	15338	9585
		low	301	260	14370	8732
Child purchase				94		3116
Total			2870	2289	161,392	99,290

Notes: Households with "children" are households with at least one person aged below 18, "Pensioners" refers to a households with no more than two people, no-one aged below 18 and at least one person aged above 64; "No children" refers to all other households. "Child purchase" refers to someone aged below 18 making a food on-the-go purchase. Skill levels are defined using socioeconomic groups. "High" comprises people in managerial, supervisory or professional roles, "low" refers to both skilled and unskilled manual workers and those who depend on the state for their income. Income levels are defined by terciles of the within household type income per person distribution. The total number of households and individuals is less than the sum of the number in each category because households may switch group over time.

Advertising data

We use advertising data collected by AC Nielsen. We have information on advertising expenditure by brand and month over the period 2001-2010. The information on earlier periods allows us to compute advertising stocks, taking into account a long period of prior advertising flows. For each brand we observe total monthly advertising expenditure, including expenditure on advertising appearing on TV, in press, on radio, on outside posters and on the internet. Advertising is at the brand level, it does not vary by pack size.

Table 2.4 describes monthly advertising expenditure. Walkers spends the most on advertising. The most advertised brand is Walkers Regular, with on average £500,000 expenditure per month. Walkers Regular also has the largest market share. The table shows the minimum and maximum advertising expenditures by month over June 2009 - October 2010. Advertising expenditures vary a lot across brands, but also across months within a brand. All brands have some periods of zero advertising expenditure, and some brands effectively never advertise, meaning that for these brands the stock of advertising is always very low.

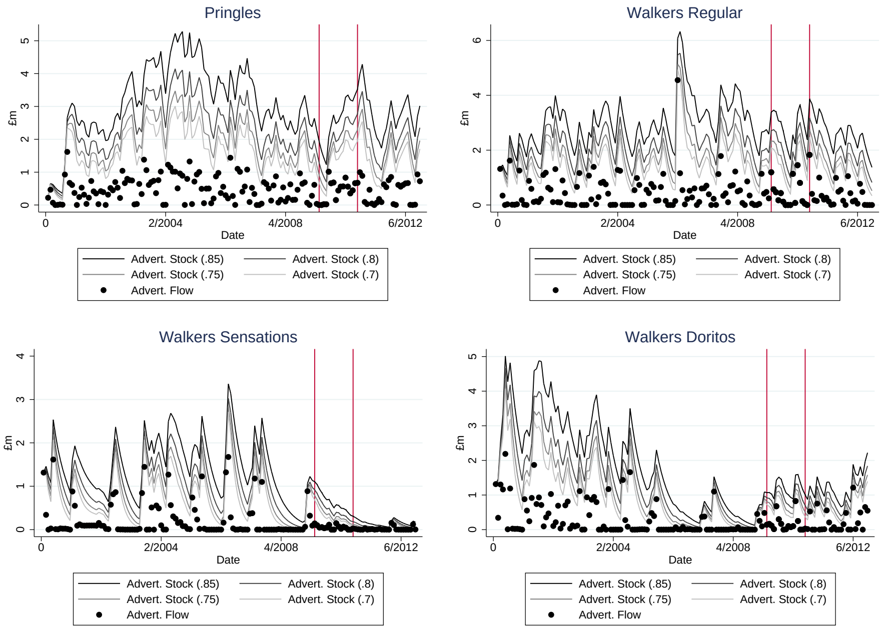
Table 2.4: *Advertising expenditures*

	Monthly expenditure (£100,000)			Total
	Mean	Min	Max	(06/09-10/10)
Pringles	4.50	0.00	10.14	76.54
Walkers Regular	4.97	0.00	18.29	84.47
Walkers Sensations	0.54	0.00	1.46	9.12
Walkers Doritos	1.75	0.00	8.25	29.67
Walkers Other	2.89	0.00	8.99	49.07
KP	2.09	0.00	8.49	35.60
Golden Wonder	0.08	0.00	0.80	1.34
Asda	0.01	0.00	0.23	0.23
Tesco	0.08	0.00	0.68	1.44
Other	1.58	0.00	5.74	26.83

Notes: Expenditure is reported in £100,000 and includes all expenditure on advertising appearing on TV, in press, on radio, on outside posters and on the internet.

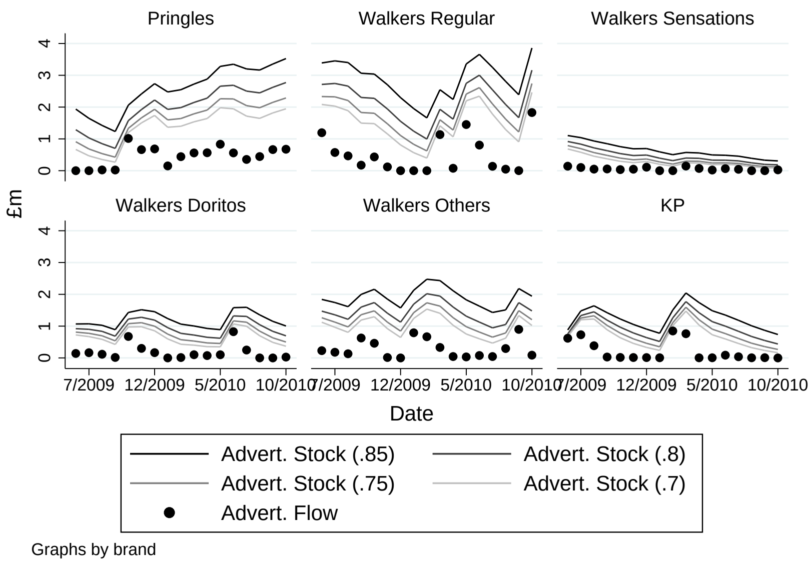
Figure 2.1 shows the monthly advertising expenditures flows for the brand Pringles and for three of the brands produced by the largest firm Walkers over the period 2000-2012. It shows that within brands, expenditures are volatile - moving between zero and relatively large values. The figure also plots the evolution of advertising stocks constructed under values of δ equal to 0.7, 0.75, 0.8 or 0.85. Stocks computed using the different decay parameters all co-move closely. Figure 2.2 shows advertising expenditures and stocks for the six most advertised brands in the market, focusing on the period June 2009 to October 2010 over which we estimate demand. The close co-movement in stocks for the different values of δ means that demand estimates are insensitive to the different values: the principle effect of increasing δ from 0.7 through to 0.9 is to change the magnitudes of the coefficients on advertising close to proportionately. As we increase δ to values close to 1, advertising stocks becomes very large and vary little over 2009-2010 because the historical stock is of a much larger order of magnitude than advertising expenditure flows. This results in very imprecise advertising effects in demand as stocks become close to collinear with the brand effect. However, if we estimate demand including advertising flow and lagged advertising stock (with δ close 1), the flow advertising has a significant effect on demand, ruling out very high values of δ. On the other hand, if we include advertising flow in demand and a lagged advertising stock variable (computed with δ sufficiently far from 1) the lagged stock variables is statistically significant, ruling out $\delta = 0$. In estimation we fix $\delta = 0.75$ - using any other value (except low and very high levels for which these is evidence against) yields very similar demand predictions.

Figure 2.1: *Advertising flows, stocks for main brands over 2000-2012*



Note: Each graph shows advertising flow expenditures in dots, advertising stocks (solid lines) for different values of δ, for the 2000-2012 period. The data used for estimation of demand are between the two vertical lines.

Figure 2.2: *Advertising flows, stocks for main brands over 2009-2010*



Note: Each graph shows advertising flow expenditures in dots, advertising stocks (solid lines) for different values of δ, for the period June 2009 to October 2010 used in demand estimation.

Figure 2.1 also shows that, although stocks depend on all lagged advertising flow expenditures since 2000, in some periods (included over the period we estimate demand) they can be close to zero. For instance, with $\delta = 0.75$, a brand not advertising for 6 months will have a stock less that 20% the value of the stock 6 months previous. A consequence of lumpy advertising is that in our counterfactual simulation of an advertising ban, where we simulate zero stocks, we are not predicting very far out of the observed sample. It also highlights that the long term advertising ban equilibrium with very close to zero advertising stocks would be reached after less than a year of transition.

Advertising in the UK potato chips market consists mainly of celebrities endorsing brands. The typical adverts show a sports star or a model eating potato chips. Prominent examples are shown in Figure 2.3. The advertisement on the top left shows supermodel Elle Macpherson eating Walkers potato chips; the one on the lower left shows an ex-professional football player and TV personality Gary Lineker with the FA Cup (football) trophy full of Walkers potato chips; the top right shows one of a series of adverts for KP Hola Hoops aimed at children, and the bottom right shows a model with Golden Wonder Skins. Given the adverts are for well established brands, and they convey little direct information content (e.g. like information on prices), it seems unlikely to us that advertising is playing an informative role in this market (this statement is true also of other junk food markets like soft drinks). For interpretation of our results on the impact of the advertising ban on demand, prices and nutrient purchases, we can remain agnostic about the role advertising plays (characteristic, persuasive of informative). However, in Section 2.5, where we consider the impact of the ban on a monetary measure of consumer welfare, we must be most specific.

Figure 2.3: *Example adverts for potato chip brands*

Note: Adverts are for Walkers (upper left), KP Hula Hoops (upper right), Walkers (lower left) and Golden Wonder Skins (lower right).

2.4.2 Empirical estimates

We estimate the demand model outlined in Section 2.2 using simulated maximum likelihood, allowing all parameters to vary by demographic groups (defined in Table 2.3) and by whether the purchase occasion is for consumption at home or on-the-go. We include random coefficients on brand advertising, competitor advertising, price and on a firm dummy for Walkers in the food at home segment. All random coefficients are assumed to have normal distributions, except those on price, which are assumed to be log normal.

We report the full set of estimated coefficients, along with the market own and cross price elasticities and marginal cost estimates in Appendix Section A.4. Here we focus on what the estimates imply for how advertising affects consumer demand. We show the impact of advertising on consumers' willingness to pay for the nutrient characteristic, price elasticities and patterns of cross brand and cross pack size substitution.

We compute the willingness to pay for a one point reduction in the nutrient profiling score (which corresponds to an increase in product healthiness), details of this calculation are in Appendix Section A.2. Table 2.5 shows the median willingness to pay across households for food at home purchase occasions and across individuals for food on-the-go purchase occasions; 95%

confidence intervals are given in brackets.[7] We evaluate the consumers' willingness to pay at three levels of advertising: zero, medium (corresponding to the average stock of the brand KP), and high (corresponding to the average stock of the brand Walkers Regular). When there is no advertising, households are willing to pay 4.8 pence per transaction for a one point reduction in the nutrient profiling score for food at home; this falls to 3.2 pence when advertising is at a medium level, and falls further to 0.4 (which is not statistically different from 0) when advertising is at a high level. Expressed as a percentage of the mean price of potato chips available for food at home purchases, households are willing to pay an additional 2.3% for a 1 point reduction in the nutrient profiling score in the absence of advertising, this falls to 0 when advertising is high. A similar pattern holds for food on-the-go, with willingness to pay for a one point reduction in the nutrient profiling score falling from 1.6% of mean price to -1.5% as advertising changes from zero to high. Table 2.5 makes clear that one thing that advertising does is lower consumers' willingness to pay for an increase in the healthiness of potato chips.

Table 2.5: *Effect of advertising on willingness to pay for 1 point reduction in nutrient profiling score*

		Advertising level		
		None	Medium	High
Food in the home	Willingness to pay in pence	4.8	3.2	0.4
		[4.1, 5.3]	[2.7, 3.6]	[-0.5, 1.7]
	% of mean price	2.3	1.6	0.2
		[2.0, 2.5]	[1.3, 1.7]	[-0.2, 0.8]
Food on-the-go	Willingness to pay in pence	0.8	-0.1	-0.7
		[0.6, 1.0]	[-0.2, 0.1]	[-0.9, -0.6]
	% of mean price	1.6	-0.2	-1.5
		[1.2, 2.0]	[-0.4, 0.2]	[-1.8, -1.1]

Notes: Numbers in rows 1 and 3 are the median willingness to pay in pence for a one point reduction in the nutrient profiling score. Numbers in rows 2 and 4 are the willingness to pay expressed as a percentage of the mean price of potato chips on the purchase occasion (i.e. food at home or food on-the-go occasion). Medium advertising refers to the mean advertising stock of the brand KP. High advertising refers to the mean advertising stock of the brand Walkers Regular. 95% confidence intervals are given in square brackets.

In our demand specification we allow advertising to interact with the price coefficient, meaning it can potentially shift consumers' price sensitivities. We find that, for the food at home segment (which represents 86% of the market) advertising leads to a reduction in consumers' sensitivity to price. In order to illustrate the strength of this effect we do the following. For each of the food at home products belonging to the three most highly advertised brands, we report, in Table 2.6, the mean market own price elasticity at observed advertising levels and the elasticity

[7] We calculate confidence intervals in the following way. We obtain the variance-covariance matrix for the parameter vector estimates using standard asymptotic results. We then take 1000 draws of the parameter vector from the joint normal asymptotic distribution of the parameters and, for each draw, compute the statistic of interest, using the resulting distribution across draws to compute Monte Carlo confidence intervals (which need not be symmetric around the statistical estimates).

if the brand was not advertised in that month (and all other brands advertising had remained at observed levels). Table 2.6 shows that the mean market own price elasticity at observed advertising levels for the 150-300g pack of Walkers is -1.5 and the elasticity for the 300g+ pack size is -2.2. If Walkers unilaterally stopped advertising, demand for its Regular brand, for both the 150-300g pack and the 300g+ pack, would become more elastic; the own price elasticities would be -1.6 and -2.5. A similar pattern is apparent for Pringles and (to a lesser extent) KP.

Table 2.6: *Effect of advertising on own price elasticities*

	Walkers Regular		Pringles		KP	
	Observed advertising expenditure	Zero advertising expenditure	Observed advertising expenditure	Zero advertising expenditure	Observed advertising expenditure	Zero advertising expenditure
<150g					-1.31	-1.35
					[-1.36, -1.27]	[-1.40, -1.31]
150g-300g	-1.48	-1.61	-1.38	-1.52	-1.67	-1.72
	[-1.54, -1.42]	[-1.68, -1.56]	[-1.44, -1.33]	[-1.58, -1.47]	[-1.73, -1.62]	[-1.78, -1.67]
300g+	-2.18	-2.53	-2.35	-2.72	-2.75	-2.86
	[-2.30, -2.09]	[-2.65, -2.44]	[-2.47, -2.23]	[-2.85, -2.61]	[-2.87, -2.65]	[-2.98, -2.76]

Notes: For each brand in the first row, we report the mean market own price elasticity for each pack size available in the food at home segment. We report the elasticity both at the level of advertising expenditure observed in the data, and if current market brand advertising was unilaterally set to zero. 95% confidence intervals are given in square brackets.

We undertake a similar exercise to illustrate the impact advertising has on brand demand. For each brand in turn, we simulate what market demand would have been in each market (month) if that brand had not been advertised in that month (and all other brands' advertising had remained at observed levels). In Table 2.7 we report the results for the highly advertised brands. If Walkers unilaterally stopped advertising its Regular brand quantity demanded for that brand would fall by 6%; demand for Pringles would increase by 4%, while demand for most other brands, and for potato chips overall, would fall. Unilaterally shutting down Pringles' advertising results in a large reduction in the quantity demanded of 21% for that brand, demand for Walkers Regular rises by around 1%, but demand for most other brands either is unaffected or falls. The overall effect is to reduce potato chip demand by 1%.

Table 2.7 makes clear that, for a number of brands, advertising is cooperative. The fact that we find evidence of cooperative advertising effects underlines the importance of allowing advertising to enter demand in a flexible way that does not unduly constrain the impact of advertising on demand a priori; if we had only included own brand advertising in the payoff function and omitted the interaction with other characteristics then the functional form assumptions would have ruled out cooperative advertising effects.

Table 2.7: *Effect of advertising on brand demand*

	Walkers Regular	Pringles	KP
Advertising expenditure (£m)	0.497	0.450	0.209
% change in brand demand if advertising expenditure is set to zero			
Walkers Regular	-5.73	1.46	0.68
	[-6.83, -4.76]	[1.17, 1.83]	[0.56, 0.81]
Walkers Sensations	-1.66	-0.55	-0.44
	[-2.03, -1.30]	[-0.89, -0.19]	[-0.56, -0.33]
Walkers Doritos	-0.80	0.03	-0.34
	[-1.26, -0.41]	[-0.33, 0.41]	[-0.50, -0.19]
Walkers Other	1.34	0.69	0.52
	[0.94, 1.73]	[0.36, 1.05]	[0.38, 0.66]
Pringles	4.09	-21.49	0.30
	[3.49, 4.80]	[-23.15, -19.91]	[0.15, 0.44]
KP	0.06	0.04	-3.29
	[-0.33, 0.45]	[-0.34, 0.45]	[-3.96, -2.75]
Golden Wonder	-3.79	-1.02	-1.54
	[-4.22, -3.36]	[-1.40, -0.59]	[-1.73, -1.36]
Asda	-2.18	-1.27	-0.95
	[-2.59, -1.74]	[-1.66, -0.84]	[-1.13, -0.76]
Tesco	-2.07	-1.13	-0.86
	[-2.52, -1.62]	[-1.55, -0.69]	[-1.04, -0.66]
Other	-0.96	-0.14	-0.40
	[-1.36, -0.59]	[-0.55, 0.27]	[-0.54, -0.25]
% change in total potato chips demand if advertising expenditure is set to zero			
	-1.64	-1.23	-0.51
	[-1.95, -1.39]	[-1.54, -0.91]	[-0.63, -0.39]

Notes: For each brand in the first row, in each market, we unilaterally set current brand advertising expenditure to zero. Numbers in the table report the resulting percentage change in quantity demanded for all brands and for the potato chips market as a whole. Numbers are means across markets. 95% confidence intervals are given in square brackets.

Table 2.8 shows how setting market advertising to zero for each of the most advertised brands affects demand for each of the pack sizes available for food at home. For Walkers Regular and KP it is demand for the largest pack size that declines when advertising expenditure is set to zero; demand for the smaller pack sizes actually increases (although by less than the fall in demand for the larger pack sizes). For Pringles, demand for both pack sizes falls, but the reduction in demand is much larger for the largest pack size. This highlights that advertising more of a particular brand leads consumers to switch to the larger pack size of the brand.

Table 2.8: *Effect of advertising on demand by pack size*

	Walkers Regular	Pringles	KP
Advertising expenditure (£m)	0.497	0.450	0.209
Change in own brand demand by pack size in 1,000kg if advertising expenditure is set to zero			
<150g			1.71
			[1.29, 2.12]
150g-300g	34.44	-6.99	5.39
	[28.44, 39.70]	[-9.44, -4.59]	[2.34, 8.24]
300g+	-257.67	-194.84	-76.49
	[-299.86, -217.90]	[-212.47, -179.21]	[-88.83, -65.04]
Change in own food at home brand demand in 1,000kg if advertising expenditure is set to zero			
	-223.23	-201.84	-69.39
	[-267.46, -181.89]	[-221.53, -184.86]	[-84.62, -55.91]

Notes: For each brand in the first row, in each market, we unilaterally set current brand advertising expenditure to zero. Numbers in the table report the change in quantity demands for all pack sizes of the brand available on food at home purchase occasions. Numbers are means across markets. 95% confidence intervals are given in square brackets.

2.4.3 Counterfactual analysis of advertising ban

We compare the observed market equilibria with one in which the advertising stocks of all firms are set to zero (i.e. to the situation after advertising has been banned for long enough for the stock to fully depreciate, which with the value $\delta = 0.75$ takes less than a year to have stocks being only 3% of their value prior to when advertising is banned). We find the new equilibrium in all markets (months) and report the means across markets.

Impact on Market Equilibrium

One effect advertising has on consumer demand is to lower consumers' sensitivity to price (see Table 2.6). Banning advertising therefore leads to toughening price competition. The (quantity weighted) average price in the market falls by 18%. This fall is driven by price reductions for products in the food at home segment that belong to the most heavily advertised brands. Table 2.9 shows the mean market price in the observed equilibrium with advertising and in the counterfactual equilibrium in which advertising is banned for the food at home products belonging to the three most advertised brands. The ban results in a fall in price for each product in Table 2.9. Walkers reduces the price of its most popular brand by the most, reducing the price of the 150-300g pack by 38p (or 30%) and the 300g+ pack by 60p (or 22%). The price of other products available in the food at home segment, belonging to brands that have lower levels of advertising, fall by less, or not at all. Prices in the smaller food on-the-go segment actually increase slightly in response to the advertising ban.

Table 2.9: *Effect of advertising ban on equilibrium prices*

	Walkers Regular		Pringles		KP	
	Pre ban equilibrium	Advertising banned	Pre ban equilibrium	Advertising banned	Pre ban equilibrium	Advertising banned
<150g					0.86	0.75
						[0.74, 0.76]
150g-300g	1.26	0.88	1.11	0.86	1.19	1.07
		[0.85, 1.58]		[0.83, 0.88]		[1.06, 1.08]
300g+	2.79	2.18	2.61	2.18	2.39	2.22
		[2.13, 2.80]		[2.13, 2.22]		[2.21, 2.24]

Notes: Numbers show the mean price across markets in £s. "Observed equilibrium" refers to the prices observed in the data; "Advertising banned" refers to counterfactual prices when advertising is banned. 95% confidence intervals are given in square brackets.

Table 2.10 summarizes the overall impact of an advertising ban on total monthly expenditure on potato chips and the total quantity of potato chips sold.[8] It also shows the impact of the ban on the mean probability a household buys potato chips on a purchase occasion and the average pack size of potato chips purchased, conditional on choosing an inside option. The first column shows the average of each variable across markets in the observed pre-ban equilibria, the second column shows numbers in the counterfactual when advertising is banned but prices are held constant, and the final column shows the numbers for new equilibria when advertising is banned and firms reoptimize prices.

[8] To gross the numbers up from our sample to the UK market we need a measure of the total market size M_t and how it is split between food at home and food on-the-go segments. From Mintel we know that total annual potato chip expenditure in the UK is around £1200m (http://www.marketingmagazine.co.uk/article/1125674/sector-insight-crisps-salty-snacks) and from the Living Cost and Food Survey we know that 14% of potato chips by volume were purchased as food on-the-go. Based on this information we can compute the implied potential market size and the size of each segment of the market.

Table 2.10: *Effect of advertising ban on purchases*

	Pre ban equilibrium	Advertising banned no price response	with price response
Expenditure (£m)	102.63	91.86	94.74
	[101.10, 103.63]	[86.50, 96.49]	[89.62, 98.76]
% change		-10.50	-7.69
		[-15.13, -5.89]	[-12.14, -3.67]
Quantity (mKg)	14.91	12.90	15.65
	[14.69, 15.06]	[12.15, 13.65]	[14.85, 16.39]
% change		-13.46	4.99
		[-18.02, -8.51]	[0.18, 10.17]
Probability of selecting potato chips	0.37	0.35	0.36
	[0.36, 0.37]	[0.33, 0.37]	[0.34, 0.38]
% change		-3.47	-1.40
		[-9.30, 1.49]	[-6.72, 3.97]
Mean pack size conditional on purchase	0.17	0.15	0.18
	[0.17, 0.17]	[0.14, 0.16]	[0.17, 0.19]
% change		-10.37	6.33
		[-14.39, -6.33]	[1.83, 11.21]

Notes: Percentage changes are shown below variables. "No price response" refers to the situation where advertising is banned and prices are held at their pre ban level; "with price response" refers to the situation where advertising is banned and firms reoptimize their prices. Expenditure refers to total expenditure on potato chip and quantity refers to the total amount of potato chips sold. Numbers are means across markets. 95% confidence intervals are given in square brackets.

In the current equilibria with advertising total monthly expenditure on potato chips was around £100m and total quantity sold was 15m kg. The impact of the ban if we hold prices constant is to induce a 10% fall in expenditure and a 13% fall in quantity sold. The reduction in quantity is driven both by households switching out of the market and buying potato chips in smaller pack sizes. When we account for the fact that oligopolistic firms will respond to the advertising ban by adjusting prices we find that expenditure falls by 8% but total quantity sold actually *increases*. The reason is that firms respond to the advertising ban by lowering prices - particularly in the food at home segment - and this leads to consumers buying more potato chips on a purchase occasion.

Impact on Health

The key motivation for advocates of advertising restrictions in junk food markets is to lower consumption of nutrients associated with diet related health problems (see for instance, WHO (2010) and Gortmaker et al. (2011)). However, whether banning advertising does reduce consumption of targeted nutrients will depend on both how advertising affects demand, including demand substitutions across goods, and on the equilibrium pricing response of firms operating in the market.

Table 2.11 summarizes the impact of the ban on purchases of nutrients. The top panel describes the impact of the ban on the total annual quantity of energy (calories), saturated fat and salt that households buy as potato chips. The bottom panel describes the impact on the nutrient content of the potato chips that households buy.

Holding prices at their pre-ban level, the advertising ban leads to a reduction in the total quantity of energy (by 13%), saturated fat (by 17%) and salt (by 14%) consumers purchase from potato chips. Conditional on purchasing potato chips, consumers also buy healthier varieties; the nutrient score of purchases falls by 3% (which corresponds to an increase in healthiness), and the quantity of saturated fat and salt per 100g of potato chip purchases fall buy 4% and 1%. Abstracting from the equilibrium response of firms, the advertising ban appears successful in improving the nutritional content of consumers' purchases.

However, as firms respond to the ban by reducing many of their prices, the health gains are to some extent over-turned once we account for the pricing response of firms. The full effect of the ban (accounting for the pricing response of firms) is actually to increase total calories by 5%; after the ban consumers face lower prices and therefore purchase more potato chips. However, the pricing response reinforces the switch consumers make to healthier varieties; the full effect of the ban is to improve the nutrient score by 5%, and to reduce the saturated fat and salt content per 100g of potato chip purchases by 8% and 3%. Therefore, even though consumers are led to purchase more potato chips they do no actually purchase a statistically significantly larger amount of saturated fat or salt.

Our assessment of the health implications of the advertising ban takes account of substitution across potato chip products and substitution away from the market. However, we do not explicitly model substitution between different potato chips and other products. Clearly, the health consequences of the ban will be affected by what are the most closely substitutable products with potato chips. If an advertising ban is implemented on potato chips alone, it is likely that there will be some substitution between potato chips and other junk food markets (for instance, the lower price in the potato chip market may induce some consumers to switch from confectionary to potato chips). However, if a ban is implemented on a broader set of junk food markets, we expect such cross junk market substitution to be more limited.

Table 2.11: *Effect of advertising ban on nutrient purchases*

	Pre ban equilibrium	Advertising banned no price response	with price response
Energy (b kj)	316.18	273.73	333.07
	[311.40, 319.48]	[257.63, 289.44]	[315.93, 348.55]
% change		-13.42	5.34
		[-17.99, -8.49]	[0.50, 10.54]
Saturates (1000 kg)	587.53	485.60	566.49
	[578.40, 593.96]	[455.91, 515.31]	[536.07, 596.35]
% change		-17.35	-3.58
		[-21.92, -12.37]	[-8.26, 1.65]
Salt (1000 kg)	266.94	229.18	270.80
	[263.07, 269.80]	[215.36, 242.74]	[257.24, 283.95]
% change		-14.15	1.45
		[-18.70, -9.27]	[-3.24, 6.40]
Nutrient score	13.81	13.43	13.12
	[13.78, 13.83]	[13.37, 13.51]	[13.05, 13.21]
% change		-2.73	-4.98
		[-3.14, -2.24]	[-5.43, -4.36]
Saturates intensity (g/100g)	3.94	3.76	3.63
	[3.92, 3.96]	[3.74, 3.80]	[3.59, 3.67]
% change		-4.48	-7.96
		[-5.14, -3.69]	[-8.74, -6.92]
Salt intensity (g/100g)	1.79	1.78	1.73
	[1.79, 1.79]	[1.77, 1.78]	[1.73, 1.74]
% change		-0.80	-3.24
		[-1.06, -0.52]	[-3.54, -2.84]

Notes: Percentage changes are shown below variables. "No price response" refers to the situation where advertising is banned and prices are held at their pre ban level; "with price response" refers to the situation where advertising is banned and firms reoptimize their prices. Nutrient score reports the mean nutrient profiling score for potato chip purchases; a reduction indicates consumers are switching to more healthy potato chips. Numbers are means across markets. 95% confidence intervals are given in square brackets.

2.5 Welfare

In this section we consider the impact of the ban on traditional economic measures of welfare. In particular, we report the impact of the ban on consumer welfare and profits of firms that manufacture and sell potato chips.[9] To understand the effect of the ban on consumer welfare it is necessary to take a view on how advertising affects consumers' underlying utility (is it informative about product characteristics, persuasive or a characteristic). In Section 2.5.1 we outline how we can capture consumer welfare in the two most plausible cases in junk food markets - persuasive and characteristic advertising. In Section 2.5.2 we the report the welfare numbers.

Note that our welfare measure does not take into account any long run health consequences that results from the ban that are not taken into account by consumers at the point of purchase.

[9]Profits of firms in the advertising industry may also be effected. Though we have less to say about this, we can state the total advertising budgets, which represent an upper bound on advertisers' profits.

However, the numbers in Table 2.11 can be combined with estimates from the medical literature to say something about monetary consequences of long term health effects.

2.5.1 Measuring consumer welfare

Our aim in specifying the demand model presented in Section 2.2 is to ensure the specification is flexible enough to capture the impact of pricing and advertising on demand regardless of which view (informative about product characteristics, persuasive or a characteristic) we take about advertising. However, to understand how a ban on advertising will affect consumer welfare requires that we take a stance on which view of advertising is most appropriate.

Advertising in the UK potato chips market principally consists of celebrity endorsements (see Section 2.4.1). This is true of much advertising in consumer goods markets, and particularly in junk food markets. We firstly consider the view that advertising is persuasive and acts to distort consumer decision making. The persuasive view of advertising has a long tradition in the advertising literature (Robinson (1933), Kaldor (1950)). More recently, the behavioral economics literature (see Bernheim and Rangel (2005)) has suggested advertising might lead consumers to act as non-standard decision makers; advertising providing environmental "cues" to consumers. While policies that improve cognitive processes are potentially welfare enhancing if the environmental cues have information content, persuasive advertising might distort choices in ways that do not enhance welfare. Bernheim and Rangel (2009) argue that *"choices made in the presence of those cues are therefore predicated on improperly processed information, and welfare evaluations should be guided by choices made under other conditions."* The welfare implications of restricting advertising that acts to distort decision making has been explored by Glaeser and Ujhelyi (2010), who are particularly concerned with firm advertising (or misinformation in their term) in food markets, while Mullainathan et al. (2012) consider the broad policy framework in public finance applications when consumers makes decisions inconsistent with their underlying welfare.

As pointed out by Dixit and Norman (1978), the welfare effects of changes in advertising will depend on whether one uses pre or post advertising tastes to evaluate welfare. When assessing the welfare implications of banning persuasive advertising it is natural to assess welfare changes using undistorted preferences (i.e. the parameters in the consumer's payoff function in the absence of advertising). This mirrors the distinction made by Kahneman et al. (1997) between decision and experience utility; in their terms, advertising affects choice and therefore decision utility, but it does not affect underlying or experience utility.

Under the persuasive view of advertising, decisions made when advertising is non-zero maximize a payoff function that does not coincide with the consumer's utility function. Consumers will choose the product that provides them with the highest payoff $\bar{v}_{ibst}$ as in equation (2.1), but the true underlying utility is based on the consumer's product valuation in the absence of advertising.

$$\widehat{v}_{ibst} = \alpha_{0i} p_{bst} + \psi_{0d} x_b + \eta_d z_{bs} + \xi_{ib} + \epsilon_{ibst} \tag{2.11}$$

In this case the consumer's expected utility at the advertising state and price vectors $(\mathbf{a_t}, \mathbf{p}_t)$ is given by evaluating the choice made by maximizing the payoff function (2.1) at preferences described by equation (2.11):

$$\widehat{W}_i \left(\mathbf{a_t}, \mathbf{p_t}\right) = \mathbb{E}_\epsilon \left[\widehat{v}_{ib^* s^* t}\right].$$

where we define $(b^*, s^*) = \underset{(b,s)\in\Omega_\kappa}{\arg\max}\{\bar{v}_{ibst}\}$. In this case, following the terminology of Kahneman et al. (1997), $\widehat{v}$ is the experience utility while $\bar{v}$ is the decision utility of the consumer. Noting that

$$\widehat{v}_{ibst} = \bar{v}_{ibst} - \left[\lambda_i a_{bt} + \rho_i \left(\sum\nolimits_{l\neq b} a_{lt}\right) + \alpha_{1d} a_{bt} p_{bst} + \psi_{1d} a_{bt} x_b\right],$$

we can write $\widehat{W}_i \left(\mathbf{a_t}, \mathbf{p}_t\right)$ as:

$$\begin{aligned}
\widehat{W}_i(\mathbf{a_t}, \mathbf{p_t}) =& \mathbb{E}_\epsilon[\bar{v}_{ib^* s^* t}] - \mathbb{E}_\epsilon \left[\lambda_i a_{b^* t} + \rho_i \left(\sum\nolimits_{l\neq b^*} a_{lt}\right) + \alpha_{1d} a_{b^* t} p_{b^* s^* t} + \psi_{1d} a_{b^* t} x_{b^*}\right] \\
=& \mathbb{E}_\epsilon \left[\max_{(b,s)\in\Omega_\kappa} \bar{v}_{ibst}\right] - \mathbb{E}_\epsilon \left[\lambda_i a_{b^* t} + \rho_i \left(\sum\nolimits_{l\neq b^*} a_{lt}\right) + \alpha_{1d} a_{b^* t} p_{b^* s^* t} + \psi_{1d} a_{b^* t} x_{b^*}\right] \\
=& W_i(\mathbf{a_t}, \mathbf{p_t}) - \sum_{(b,s)\in\Omega_\kappa} s_{ibst} \left[\lambda_i a_{bt} + \rho_i \left(\sum\nolimits_{l\neq b} a_{lt}\right) + \alpha_{1d} a_{bt} p_{bst} + \psi_{1d} a_{bt} x_b\right],
\end{aligned}$$

where s_{ibst} is given by equation (2.5) and, up to an additive constant,

$$\begin{aligned}
W_i \left(\mathbf{a_t}, \mathbf{p_t}\right) \equiv& E_\epsilon \left[\max_{(b,s)\in\Omega_\kappa} \bar{v}_{ibst}\right] \\
=& \ln \Bigg[\exp\left(\zeta_{d0t}\right) + \sum_{(b,s)\in\Omega_\kappa} \exp \Big(\alpha_{0i} p_{bst} + \psi_{0d} x_b + \eta_d z_{bs} + \xi_{ib} \\
& + \left[\lambda_i a_{bt} + \rho_i \left(\sum\nolimits_{l\neq b} a_{lt}\right) + \alpha_{1d} a_{bt} p_{bst} + \psi_{1d} a_{bt} x_b\right]\Big)\Bigg]
\end{aligned}$$

using the standard closed form (Small and Rosen (1981)) when the error term ϵ is distributed type I extreme value.

This says that when a consumer's choices are distorted by advertising, expected utility is equal to expected utility if advertising was in the consumer's utility function, minus a term reflecting the fact that the consumer is making choices that do not maximize her true underlying utility function.

Denote $\mathbf{p}^0$ a price counterfactual equilibrium in which there is no advertising (defined in Section 2.3). Evaluating the impact of banning advertising under the welfare standard of $\widehat{v}_{ibst}$, the consumer welfare difference between the equilibrium with advertising and the one in which advertising is banned can be decomposed as:

$$\begin{aligned} W_i\left(\mathbf{0},\mathbf{p_t^0}\right) - \widehat{W}_i\left(\mathbf{a_t},\mathbf{p_t}\right) = & W_i\left(\mathbf{0},\mathbf{p_t}\right) - \widehat{W}_i\left(\mathbf{a_t},\mathbf{p_t}\right) \quad \text{(choice distortion effect)} \\ & + W_i\left(\mathbf{0},\mathbf{p_t^0}\right) - W_i\left(\mathbf{0},\mathbf{p_t}\right) \quad \text{(price competition effect)} \end{aligned} \tag{2.12}$$

where we use the fact that $\widehat{W}_i(\mathbf{0},\mathbf{p}) = W_i(\mathbf{0},\mathbf{p})$.

Under this persuasive view of advertising, advertising has the effect of inducing the consumer to make suboptimal choices. Banning advertising removes this distortion to decision making, which benefits consumers. We label this the "choice distortion effect". However, banning advertising also affects consumer welfare through the "price competition effect" channel. The sign of this effect will depend on the change in pricing equilibrium. The price competition effect is independent of the view we take about advertising since firms' behavior depends only on decision utilities of consumers.

An alternative to the persuasive view of advertising is that it is a characteristic of the product that consumers value (Stigler and Becker (1977) and Becker and Murphy (1993)). In this case, in the terminology of Kahneman et al. (1997), advertising would enter both experience and decision utilities. The welfare effect of banning advertising would be given by the more standard term $W_i\left(\mathbf{0},\mathbf{p_t^0}\right) - W_i\left(\mathbf{a_t},\mathbf{p_t}\right)$ and the choice distortion term in the equation (2.12) would be replaced by a term reflecting the impact on welfare of removing the advertising characteristic from the market, $W_i\left(\mathbf{0},\mathbf{p_t}\right) - W_i\left(\mathbf{a_t},\mathbf{p_t}\right)$.

The identification of this characteristic effect is influenced by the normalization of the outside option utility. We include own brand and competitor advertising in the payoff function of inside goods, but the alternative specification where own brand advertising appears in the payoff of inside goods and total advertising appears in the payoff of the outside option would give rise to observationally equivalent demand. Although observationally equivalent, these two specifications would lead to different welfare predictions under the characteristics view.[10] However, as advertising does not enter the experience utility under the persuasive view, this problem does not exist in this alternative welfare definition.

We focus on welfare measures of the direct monetary costs for consumers and firms of an advertising ban. To measure the monetary costs to a consumer we convert the welfare changes to compensating variation (dividing by the marginal utility of income):

$$CV_i\left(\mathbf{a_t},\mathbf{p_t},\mathbf{p_t^0}\right) = \frac{1}{\alpha_{0i}}\left[W_i\left(\mathbf{0},\mathbf{p_t^0}\right) - \widehat{W}_i\left(\mathbf{a_t},\mathbf{p_t}\right)\right] \tag{2.13}$$

[10]See Appendix Section A.3 for details.

Aggregate compensating variation is given by integrating across the observable and unobservable consumer level heterogeneity;

$$CV\left(\mathbf{a_t},\mathbf{p_t},\mathbf{p_t^0}\right) = \int CV_i\left(\mathbf{a_t},\mathbf{p_t},\mathbf{p_t^0}\right)dF(\pi_i^u,\pi_i^o,d). \tag{2.14}$$

2.5.2 Welfare impact of ban

The first three lines of Table 2.12 show the impact of the ban on consumer welfare under the persuasive view of advertising. The "choice distortion effect" in the first row is a measure of the welfare gain to consumers of no longer making decisions distorted by advertising, with prices held at their pre ban level. The second row reports the gains through increased price competition after banning advertising. The third row is the sum of these two effects. The "choice distortion effect" leads to a £17 million increase in consumer welfare. The "price competition effect" raises consumer welfare by a further £9 million. Firms, on average, respond to the ban by lowering their prices and consumers benefit from paying these lower prices. In equilibrium, the ban increases total consumer welfare, under the persuasive view of advertising, by £26 million per month.

The fourth line shows the impact of the ban on firms' variable profits (net of advertising expenditure). Banning advertising does not lead to a statistically significant change in variable profits. In Appendix Section A.4.4 we show profits by firm. The dominant firm, Walkers, plays an important role; it sells more after advertising is banned and this leads to an increase in variable profits (net of advertising expenditure). The profits of other firms fall.

Therefore under the persuasive view of advertising, the effect of the ban is to raise total welfare by £26 million; around one-third of this is due to increased price competition, and around two-thirds to removal of the distorting effects of persuasive advertising.

Table 2.12: *Effect of advertising ban on welfare under persuasive view of advertising*

	Advertising banned	
	no price response	with price response
Choice distortion effect (£m)	17.4	17.4
	[16.8, 19.6]	[16.8, 19.6]
Price competition effect (£m)	0.0	8.9
		[7.7, 9.9]
Total compensating variation (£m)	17.4	26.4
	[16.8, 19.6]	[25.4, 28.6]
Change in profits (£m)	-0.4	-0.2
	[-2.9, 2.2]	[-2.4, 2.0]
Total change in welfare (£m)	17.1	26.2
	[15.0, 20.5]	[24.0, 29.4]

Notes: "No firm response" refers to case of an advertising ban when prices are held at their pre ban level; "Firm response" refers to case of an advertising ban when firms reoptimize their prices. Compensating variation is computed under the view that advertising distorts consumer decision making. 95% confidence intervals are given in square brackets.

An alternative to the view that advertising distorts consumer decision making is the view that it affects utility directly as a characteristic that consumers value. Table 2.13 describes the changes in welfare under this view of advertising. In this case the "choice distortion effect" is replaced by the "characteristic effect". The characteristic effect is influenced by the normalization of the outside option utility. Under our adopted normalization, where own brand and competitor advertising enter the payoff function of inside goods, the characteristic effect leads to a reduction in consumer welfare of £18 million. This fall out-weighs the price competition effect, meaning total welfare falls.

Table 2.13: *Effect of advertising ban on welfare under characteristic view of advertising*

	Advertising banned	
	no price response	with price response
Characteristic effect (£m)	-18.4	-18.4
	[-22.1, -14.3]	[-22.1, -14.3]
Price competition effect (£m)	0.0	8.9
		[7.7, 9.9]
Total compensating variation (£m)	-18.4	-9.5
	[-22.1, -14.3]	[-13.3, -4.8]
Change in profits (£m)	-0.4	-0.2
	[-2.9, 2.2]	[-2.4, 2.0]
Total change in welfare (£m)	-18.8	-9.6
	[-24.9, -12.3]	[-16.0, -3.1]

Notes: "No firm response" refers to case of an advertising ban when prices are held at their pre ban level; "Firm response" refers to case of an advertising ban when firms reoptimize their prices. Compensating variation is computed under the characteristic view of advertising. 95% confidence intervals are given in square brackets.

2.6 Robustness

In this section we test the robustness of our results to various modifications to the specification. Rather than repeat all of the results tables, we focus on the impact that a firm changing it's level of advertising has on the shape of demand. We show that in all specifications advertising has a very similar impact on market demand curves.

In Section 2.2.4 we argued that our very rich data, coupled with specific features of the UK market (e.g. national pricing and advertising) mitigate many common endogeneity concerns. Nevertheless, concern may remain that our estimates are contaminated by endogeneity. Our first robustness check is therefore to repeat our analysis implementing a control function approach (see Blundell and Powell (2004) and for multinomial discrete choice models Petrin and Train (2010)). We estimate a control function for both advertising and price.

Our advertising state vector a_{bt} depends on the advertising flow e_{bt} in addition to past advertising flows e_{bt-1}, e_{bt-2}, ... Therefore, if the state variable a_{bt} is correlated with demand shocks it is most likely to be through its dependence on the contemporaneous advertising flow. We estimate a first stage regression of product level monthly advertising flows on time effects, exogenous brand characteristics (included in the demand model), and instruments chosen to be correlated with potato chip brand advertising flows but not with demand shocks. A natural instrument would be the price of advertising, unfortunately we do not directly observe advertising prices. Instead we use advertising expenditure in another market (the ready-meal sector), where shocks to demand are likely to be uncorrelated with shocks in the potato chips market. We interact the instrument with potato chip brand fixed effects. The logic is that ready-meal advertising will be correlated with potato chip advertising through the common influence of the price of advertising but be independent of ϵ_{ibst}. The first stage estimates show that the instrument has some power (a test of the joint significance of the ready-meal brand interactions has a p value less than 0.001).

Prices in the UK potato chip market are set nationally, and over the time period of our data there is very little variation in product attributes, products or sets of competitors. Therefore, as an instrument for product price, we elect to use past prices. This instrument will eliminate the influence of any market (month) specific demand shocks that are uncorrelated over time (for instance, the worry that a sale - which typically last for less than a month - is driven by a contemporaneous shock to demand). Unsurprisingly, our price instruments are highly correlated with price (conditional on exogenous variables included in the demand model).

The second robustness test we consider is to estimate an alternative specification that includes an additional measure of consumer advertising exposure in the model. For households in the food in sample we observe the amount of TV the main shopper watches over the course of a week. As the vast majority of potato chip advertising is done on national TV, information on how regularly households watch TV may help capture any residual differences in exposure to advertising not captured by the numerous consumer type specific random coefficients already included in the model. To assess this we include the measure of TV watching behavior in the model interacted with each advertising variable - the own and competitor advertising level effects and advertising-price and advertising-nutrient characteristics interactions - and we allow all these additional variables to be demographic group specific.

The third alternative specification we consider involves replacing the single index nutrient score with individual measures of saturated fat and salt. This allows for the possibility that consumers place different weights in their payoff functions on the individual nutrients than the weights implicit in the calculation of the single index.

In Table 2.14 we summarize the effect on demands that would result if Walkers Regular and Pringles, separately and unilaterally, ceased advertising in one market (month). We report the average effect across all markets. The top panel shows percent changes in brand level demands. The second panel shows the percentage change in demand for the smaller and larger food at home pack sizes for the brand that has zero advertising flow expenditure (i.e for Walkers Regular and for Pringles). The final panel reports the percent change in the absolute value of own price elasticities that would results for the two pack sizes for the brand that has zero advertising flow expenditure. We show the results for our main specification (replicating the information in Tables 2.6, 2.7 and 2.8), and for the three alternative specifications. The numbers makes clear that our findings that advertising is partially predatory and partially cooperative, it leads consumers to switch to larger pack sizes and it acts to make demand more price elastic hold across all specifications. The alternative specifications, like our baseline model, also all predict that an advertising ban at constant prices leads to a reduction in energy, saturated fat and salt purchases, but that firms respond to the ban by lowering their prices, which counteracts these falls. Although the precise size of the effects changes a little across each specification our main conclusions are unaffected.

Table 2.14: *Effect of advertising on demand: alternative specifications*

	Main specification		Control function		TV exposure		Separate nutrients	
	Walkers Regular	Pringles	Walkers Regular	Pringles	Walkers Regular	Pringles	Walkers Regular	Pringles
% change in brand demand if advertising expenditure is set to zero								
Walkers Regular	-5.73	1.46	-3.56	1.57	-5.33	1.06	-5.74	1.33
	[-6.83, -4.76]	[1.17, 1.83]	[-5.00, -2.02]	[1.24, 1.90]	[-6.42, -4.37]	[0.73, 1.35]	[-6.87, -4.69]	[1.03, 1.69]
Walkers Sensations	-1.66	-0.55	-2.22	-0.43	-1.88	-0.88	-1.40	-0.49
	[-2.03, -1.30]	[-0.89, -0.19]	[-2.66, -1.76]	[-0.81, -0.08]	[-2.24, -1.50]	[-1.26, -0.53]	[-1.77, -1.03]	[-0.83, -0.14]
Walkers Doritos	-0.80	0.03	-1.42	0.13	-1.08	-0.29	-0.63	0.03
	[-1.26, -0.41]	[-0.33, 0.41]	[-1.94, -0.88]	[-0.23, 0.49]	[-1.50, -0.68]	[-0.66, 0.04]	[-1.06, -0.18]	[-0.31, 0.39]
Walkers Other	1.34	0.69	0.77	0.77	1.00	0.33	1.33	0.63
	[0.94, 1.73]	[0.36, 1.05]	[0.27, 1.26]	[0.43, 1.09]	[0.60, 1.39]	[0.00, 0.64]	[0.91, 1.77]	[0.32, 0.96]
Pringles	4.09	-21.49	3.87	-20.81	3.15	-20.06	3.77	-20.76
	[3.49, 4.80]	[-23.15, -19.91]	[3.19, 4.60]	[-22.53, -19.11]	[2.57, 3.85]	[-21.70, -18.20]	[3.21, 4.47]	[-23.01, -18.70]
KP	0.06	0.04	-0.36	0.17	-0.31	-0.35	0.03	0.00
	[-0.33, 0.45]	[-0.34, 0.45]	[-0.81, 0.09]	[-0.24, 0.53]	[-0.69, 0.05]	[-0.74, 0.02]	[-0.37, 0.47]	[-0.37, 0.42]
Golden Wonder	-3.79	-1.02	-4.36	-0.86	-3.95	-1.38	-3.58	-1.02
	[-4.22, -3.36]	[-1.40, -0.59]	[-4.86, -3.88]	[-1.28, -0.43]	[-4.45, -3.52]	[-1.78, -1.00]	[-4.01, -3.11]	[-1.41, -0.57]
Asda	-2.18	-1.27	-1.26	-0.01	-2.34	-1.47	-2.12	-1.28
	[-2.59, -1.74]	[-1.66, -0.84]	[-1.66, -0.82]	[-0.40, 0.41]	[-2.76, -1.97]	[-1.89, -1.07]	[-2.52, -1.66]	[-1.67, -0.83]
Tesco	-2.07	-1.13	-2.27	-1.00	-2.49	-1.63	-1.99	-1.13
	[-2.52, -1.62]	[-1.55, -0.69]	[-2.72, -1.84]	[-1.43, -0.57]	[-2.87, -2.07]	[-2.01, -1.24]	[-2.41, -1.54]	[-1.51, -0.69]
Other	0.96	-0.14	-2.33	-1.12	-1.33	-0.59	-0.90	-0.16
	[-1.36, -0.59]	[-0.55, 0.27]	[-2.74, -1.88]	[-1.53, -0.71]	[-1.70, -0.95]	[-0.98, -0.20]	[-1.29, -0.45]	[-0.54, 0.30]
% change in own pack size demand if advertising expenditure is set to zero								
150g-300g	11.45	-5.78	16.13	-4.69	11.13	-5.03	11.28	-4.98
	[9.49, 13.32]	[-7.69, -3.85]	[13.75, 18.96]	[-6.99, -2.49]	[9.05, 13.03]	[-7.02, -3.05]	[9.37, 13.22]	[-7.81, -2.44]
300g+	-8.04	-23.82	-5.88	-23.43	-7.52	-22.28	-7.99	-23.08
	[-9.31, -6.91]	[-25.50, -22.26]	[-7.28, -4.34]	[-25.09, -21.70]	[-8.76, -6.41]	[-24.00, -20.46]	[-9.11, -6.85]	[-25.36, -21.08]
% change in absolute values of own price elasticity if advertising expenditure is set to zero								
150g-300g	8.93	9.88	6.56	7.62	8.43	9.47	8.91	9.98
	[8.10, 9.76]	[8.85, 10.86]	[5.94, 7.18]	[6.81, 8.47]	[7.54, 9.18]	[8.50, 10.50]	[8.13, 9.72]	[8.92, 10.99]
300g+	16.13	15.79	12.16	11.84	15.43	15.12	16.25	15.98
	[15.04, 17.28]	[14.52, 17.25]	[11.35, 13.00]	[10.92, 12.95]	[14.27, 16.56]	[13.88, 16.59]	[15.11, 17.47]	[14.63, 17.47]

Notes: The first row refers to the model specification. Main specification refers to the demand specification outlined in Section 2.2. Control function refers to demand estimates when control functions for advertising and price are used. TV exposure refers to demand estimates when we include a measure of households' TV watching exposure. Separate nutrients refers to demand estimates when we include saturated fat and salt in demand in place of the nutrient score. For brands in the second row (Walkers Regular and Pringles), in each market, we unilaterally set current brand advertising expenditure to zero. Numbers in the first panel report the resulting percentage change in quantity demanded for all brands. Numbers in the second panel report the percentage change in own demands for pack sizes available on food at home purchase occasions for the brand in the second row. The third panel reports mean market own price elasticity for each pack size available in the food at home segment for the brand in the second row. Numbers are means across markets. 95% confidence intervals are given in square brackets.

2.7 Summary and conclusions

In this paper we develop a model of demand and supply in a market where firms compete over prices and advertising budgets, and where the impact of current advertising on future demand means that each firm's problem is a dynamic one. We allow advertising to impact demand in a flexible way, which allows us to understand the impact of advertising on demand while remaining agnostic about the view taken of advertising (as informative, a characteristic or persuasive), and we do not rule out a priori that advertising is cooperative and leads to market expansion or that it is predatory and possibly leads to market contraction. We apply the model to the potato chip market using novel transaction level data on purchases of food taken into the home and food bought on-the-go for immediate consumption. We find that brand advertising increases both own demand and often competitor demand, suggesting that it is, at least in part, cooperative. As well as attracting new customers, higher brand advertising also induces consumers to trade up to larger pack sizes, reduces consumers' price sensitivities and lowers consumers' willingness to pay for healthier products.

We use the structural model to simulate the impact of an advertising ban on market equilibrium. This both helps us understand the impact that advertising has on equilibrium outcomes, and given recent calls for restrictions in junk food advertising, is an interesting exercise from a policy perspective. We find that banning advertising lowers potato chip demand, as well as total purchases of calories, saturated fat and salt only if firms do not respond by changing their prices. In the more realistic scenario in which firms re-optimize prices in response to the ban, total demand for potato chips actually rises. This is because the ban increases price competition and so firms respond by lowering average prices and the increase in demand this induces more than offsets the direct fall in demand from no advertising. In this case consumers purchase more calories in the form of potato chips than before the ban, but because they also switch to healthier varieties they do not necessarily purchase more saturated fat or salt.

Ultimately we are interested in the impact of the ban on welfare. We show how to calculate the change in welfare under different assumptions about how advertising affects experience utility. In the potato chip market, as in many junk food markets, advertisements consist mainly of celebrity endorsements of well established brands. We show how to evaluate consumer welfare under the view that advertising is persuasive, acting to distort consumer decision making, leading them to take decisions that are inconsistent with their underlying preferences. Under this view of advertising the ban acts to raise consumer and total welfare. In the counterfactual equilibrium consumers no longer make distorted decisions and benefit from lower prices, while in aggregate firms do not lose profits.

In this paper our focus has been on the impact of an advertising ban on a market with a set of well established and known brands. An interesting avenue for future research would be to consider an alternative counterfactual; for instance how would firms' pricing and advertising strategies respond to the introduction of a tax. The framework we develop in this paper could potentially be used to study such a question, although solving for the set of counterfactual equilibria would present considerable challenges. In markets with a reasonable degree of product churn, entry and exit considerations may play a more prominent role than in the potato chips market. In such a case, the ex ante evaluation of an advertising ban could be extended to study the effects of a ban on industry structure. Advertising may constitute a barrier to entry, and banning advertising may facilitate entry of competitors who would not need to invest in building up large advertising stocks.[11] While in the particular market studied in the paper, this consideration is not of first-order concern, in other less mature markets it may be more important. This represents a promising direction for future research.

[11] See, for instance Doraszelski and Markovich (2007), Chamberlin (1933), Dixit (1980), Schmalensee (1983) and Fudenberg and Tirole (1984)).

Chapter 3

Income effects and the welfare consequences of tax in differentiated product oligopoly

3.1 Introduction

Random utility models are widely used to study consumer choice among differentiated products. When using such models, it is common to make strong assumptions about the marginal utility of income. Such assumptions help with model tractability, simplify analysis of counterfactual equilibria and simplify welfare calculations. It is well understood that these assumptions are restrictive. They place strong restrictions on income effects, on the curvature of demand, and hence on predictions of pass-through (see, inter alia, McFadden (1999), Herriges and Kling (1999), Weyl and Fabinger (2013) and Fabinger and Weyl (2014)). Nevertheless, it is commonly believed that for small budget share product categories, the assumption of a constant marginal utility of income is a reasonable approximation. Many such applications allow income to enter in an ad hoc form as a "preference shifter". Alternatively, for larger budget share products, it is common to include income or total expenditure in log form.

In this paper, we show that flexibly modeling income effects can be important, particularly if one is interested in the distributional effects of a policy change, even in a market in which, a priori, the expectation is that income effects will play a limited role. We allow for much more flexible forms of income effects than is common in the applied discrete choice demand literature and, to highlight the implications of flexibly incorporating income effects, we use our model to simulate the introduction of a tax and compare the implications for demand, tax pass-through and welfare with those implied by specifications standard to the literature.

The existing literature that uses logit models to capture consumer demand for differentiated goods has typically made one of two assumptions regarding the nature of income effects. Most commonly among papers focusing on product categories that comprise a small budget share, researchers have assumed that utility, conditional on selecting a given option, is linear in income (or expenditure) minus price (see, Nevo (2001), Villas-Boas (2007)). Under this assumption, income drops out of the model when comparisons are made across alternatives and income effects are therefore ruled out. To capture the cross-sectional relationship between income and purchase patterns, researchers often include income in a reduced form way as a preference shifter, which linearly shifts the coefficient on price. The second common assumption made in the literature, usually when the application relates to product categories that comprise a large share of consumers' budgets, is that conditional utility is linear in the log of income (or expenditure) minus price (for example, Berry et al. (1995), Goldberg and Verboven (2001) and Petrin (2002)). This specification incorporates income effects into the model, but does so in a restrictive way. At the consumer level, the conditional marginal utility of income is inversely proportional to income.

We show that neither of these standard models can fully replicate the results we obtain with our more general model with flexible income effects. The log utility model produces estimates of demand, pass-through and welfare that are biased, both at the average and across the total expenditure distribution. Although the log utility model does allow for income effects, the restrictive function it imposes on the conditional marginal utility of income is strongly rejected in more flexible specifications. In contrast, the preference shifter model yields estimates of market level average quantities, such as tax pass-through, average price elasticities and aggregate welfare effects, that are similar to a model with flexible income effects. However, it fails to fully recover variation in price sensitivity and welfare effects across the expenditure distribution.

The preference shifter model does not admit income effects. However, it does allow for some cross sectional correlation in demand patterns and welfare effects with total expenditure through including the latter as a linear shifter of the price coefficient. We show that when income effects exist in demand, but when the counterfactual of interest involves relatively small price changes that do not themselves induce large income effects, a simple modification to the standard preference shifter model, which involves interacting price with higher order expenditure terms, can do a very good job of replicating the distributional results found with the full model with flexible income effects. The reason for this is that, even though a consumer's utility function may be highly nonlinear, for small changes in price it can be well approximated by a linear function.

Therefore the correctly linearized model - which can be approximated by interacted price with functions of total expenditure (or income) - performs well when analyzing impacts of small price changes.

On the other hand, if one is interested in studying a policy reform that shifts prices by a large amount relative to total expenditure (for instance in a market for large budget share goods) or if one is interested in understanding how consumers would respond to a policy that changed total expenditure or income, it is necessary to estimate the model with flexible income effects to obtain unbiased estimates of the effect of the reform.

We investigate the empirical relevance of these issues in a market where income effects would not seem to be a major concern, the butter and margarine market. In the UK, this market represents only about 1% of average total grocery expenditure. Nonetheless, in common with many other product categories, consumer purchase patterns are strongly related to total grocery expenditure. In particular, higher total grocery expenditure is associated with a higher probability of purchase, and, conditional on purchase, selecting a relatively high priced option. We estimate demand in this market allowing a cubic spline to capture the structural relationship between a consumer's net expenditure and their utility from purchasing an option, and we show that the relationship is nonlinear and approximately cubic. Standard specifications are unable to recover the distributional consequences of introducing a tax, but we show the correctly linearized model is successful in doing this..

The fact that, in the case of small price changes, a linear approximation of the utility model with flexible income effects succeeds in recovering the distributional patterns across consumers is potentially useful because computing counterfactual equilibria and evaluating welfare effects of a price change in the model with income effects can be considerably more costly. In particular, in the model with income effects the simple formula for compensating variation from Small and Rosen (1981) is not valid and to compute compensating variation one must use either the simulation methods introduced in McFadden (1999) or Dagsvik and Karlström (2005). Recently Bhattacharya (2015) has shown how to estimate the marginal distribution of compensating variation non-parametrically when interest centers on the impact of a change in the price of a single good. However, in differentiated product markets in which interest typically centers on estimation of the welfare impacts of simultaneous changes in multiple prices, these methods are not applicable.

This reliance of the discrete choice literature on restrictive assumptions about the nature of income effects contrasts with the continuous choice demand literature which has concerned itself with allowing for increasingly general forms of income effects (see for instance, Deaton and Muellbauer (1980), Banks et al. (1997), Lewbel and Pendakur (2009), Hausman and Newey

(2014)). Researchers in the continuous choice demand literature have found that flexible models of income effects are important for understanding demand patterns. We find that the same is true in discrete choice models.

Assumptions about income effects in random utility models may also have a strong bearing on patterns of tax pass-through and on price increases predicted by merger simulations. A series of papers (including Seade (1985), Delipalla and Keen (1992) and Anderson et al. (2001)) provide theoretical pass-through results in stylized models of imperfect competition (with either homogenous or symmetrically differentiated goods). Weyl and Fabinger (2013) provide a framework which nests many of the previous theoretical results, and highlights the importance of a number of determinants of pass-through. All of these papers highlight the important role that the curvature of market demand plays in determining tax pass-through. Constraining the form of income effects in logit demand models restricts the curvature of individual consumer level demand curves. Market demand curves may still be somewhat more flexible if preference heterogeneity is included in the model, but they are nonetheless influenced by assumptions made about consumer level demands. We explore the importance of relaxing demand curvature restrictions through allowing for flexible income effects when assessing equilibrium pass-through of a tax to consumer prices.

Our work is related to a large literature that estimates pass-through of cost shocks and taxes to prices. A series of papers use observed tax changes to estimate pass-through. These include Besley and Rosen (1999), who exploit variation in State and local sales taxes in the US and look at the impact on prices of a number of products, Delipalla and O'Donnell (2001), who analyze the incidence of cigarette taxes in several European countries and Kenkel (2005), who uses data on how the price of alcoholic beverages changed in Alaska. Results from the literature vary, but typically these papers find complete or overshifting of excise taxes, which broadly accord with our pass-through results.

A number of papers use structural models to study equilibrium pass-through. Many of these papers find that pass-through of cost shocks is incomplete (see, for instance, Goldberg and Hellerstein (2013) and Nakamura and Zerom (2010)). An important reason for incomplete pass-through of cost shocks is that often not all cost components are affected by the shock. For instance, exchange rate movements do not directly impact the cost of non-traded inputs (Goldberg and Hellerstein (2008)). In a context where firms' marginal costs are observable (in the wholesale electricity market), Fabra and Reguant (2014) find changes in marginal costs are close to fully shifted to prices. Another feature of this literature has been to highlight that nominal rigidities may be important in generating delayed adjustment to shocks, although they are less important in determining long-run pass-through. We add to this literature by studying how equilibrium tax pass-through in an imperfectly competitive market is affected by functional form assumptions that restrict the shape of market demand.

The rest of the paper is structured as follows. In Section 3.2 we discuss various ways of modeling income effects in random utility models and their implication for measuring consumer welfare effects. In Section 3.3 we discuss market level demand and how assumptions made about consumer level demand influence the curvature of the market demand curve. Section 3.4 presents results from an empirical example. We compare how different forms of income effects in demand impact on the consumer welfare effects and pass-through of an excise tax. A final section concludes.

3.2 Consumer level demand

We consider a random utility model of consumer choice (see McFadden (1981)). The consumer has a total budget y available to spend. The variable y may be the consumer's income, or it may represent the total expenditure the consumer allocates to a set of goods over which preferences are weakly separable. For instance, in applications to a particular grocery product category, y may be total grocery expenditure. The consumer makes a discrete choice about which alternative $j \in \{0, 1, ..., J\}$ to purchase and spends their remaining budget on other groceries. We denote the price of option j as p_j. Option $j = 0$ denotes the 'outside option' and p_0=0. Option j has associated with it a vector of observable product characteristics $\mathbf{x_j}$ and unobservable characteristics ε_j. Utility from selecting option j is given by $U(y - p_j, \mathbf{x_j}, \varepsilon_j)$. We refer to $U(y - p_j, \mathbf{x_j}, \varepsilon_j)$ as the consumer's conditional utility. It is the utility obtained conditional on selecting option j. In this section, we leave implicit the dependence of U on a vector of parameters θ, some of which may be random coefficients that vary across consumers. We discuss consumer heterogeneity in more detail in Sections 3.3 and 3.4.1.

The consumer indirect utility function is given by:

$$V(\mathbf{p}, y, \mathbf{x}, \boldsymbol{\varepsilon}) = \max_{j \in \{0,...,J\}} U(y - p_j, \mathbf{x_j}, \varepsilon_j) \tag{3.1}$$

where $\mathbf{p} = (p_1, ..., p_J)'$, $\mathbf{x} = (\mathbf{x_1}, ..., \mathbf{x_J})$ and $\boldsymbol{\varepsilon} = (\varepsilon_1, ..., \varepsilon_J)'$. As long as the conditional utility function, $U(y - p_j, \mathbf{x_j}, \varepsilon_j)$, is continuous and non-decreasing in $y - p_j$, $V(\mathbf{p}, y, \mathbf{x}, \boldsymbol{\varepsilon})$ satisfies the properties of an indirect utility function; it is non-increasing in prices, non-decreasing in total budget, homogeneous of degree zero in all prices and total budget, quasi-convex in prices and continuous in prices and total budget. Consumer theory does not impose further restrictions on how $y - p_j$ enters conditional utility.

To focus on the role of income effects in the most commonly used logit model, we employ the standard assumption that $\boldsymbol{\varepsilon}$ are additive, independent and identically distributed across alternatives and drawn from a type I extreme value distribution. As shown in McFadden and Train (2000), any discrete choice model derived from random utility maximization has choice probabilities that can be approximated to any degree of accuracy by a mixed logit model. So, this restriction does not overly constrain the scope of our analysis as long as preference heterogeneity is included in the model. An alternative is to assume $\boldsymbol{\varepsilon}$ is additive and is drawn from a generalized extreme value distribution, leading, for example, to a nested logit choice model.

Under the additive assumption, an individual consumer's conditional utility is given by:

$$U(y - p_j, \mathbf{x_j}, \varepsilon_j) = \widetilde{U}(y - p_j, \mathbf{x_j}) + \varepsilon_j, \tag{3.2}$$

$$\varepsilon_j \sim \text{i.i.d. type I extreme value.}$$

The probability the consumer selects option j,

$$P_j = \Pr\left[U(y - p_j, \mathbf{x_j}, \varepsilon_j) \geq U(y - p_k, \mathbf{x_k}, \varepsilon_k) \ \forall k\right],$$

under (3.2) is given by:

$$P_j = \frac{\exp(\widetilde{U}(y - p_j, \mathbf{x_j}))}{\sum_{k \in \{0,\dots,J\}} \exp(\widetilde{U}(y - p_k, \mathbf{x_k}))}. \tag{3.3}$$

The bulk of the applied literature restricts this specification even further by imposing that the marginal utility of income is constant:

$$\widetilde{U}(y - p_j, \mathbf{x_j}) = \alpha(y - p_j) + g(\mathbf{x_j}). \tag{3.4}$$

This specification rules out income effects. At the consumer level, when comparisons are made across options, y differences out of the model. To capture the fact that choice patterns commonly vary across consumers with different budgets, it is typical to include y in the model as a "preference shifter" (see, inter alia, Nevo (2001), Berry (1994), Villas-Boas (2007)). For example, the parameter α may be allowed to vary linearly across consumers with total budget (and possibly also with other demographic variables):

$$\alpha = \alpha_0 + \alpha_1 y + \nu \tag{3.5}$$

where ν is a random coefficient. This "preference shifter" model has no income effects at the individual level and is ad hoc; consumer theory does not provide a theoretical explanation for why preferences should shift with y. However, this approach does allow researchers to capture, in a reduced form way, the empirical fact that expenditure patterns do vary cross-sectionally with income or total expenditure.

Papers that do allow for income effects include Berry et al. (1995), Goldberg and Verboven (2001) and Petrin (2002). These papers consider demand for large budget share product categories (automobiles and mini-vans) and specify

$$\widetilde{U}(y - p_j, \mathbf{x_j}) = \alpha \ln(y - p_j) + g(\mathbf{x_j}). \tag{3.6}$$

In this case the conditional marginal utility of income is given by $\frac{\alpha}{y-p_j}$; the conditional marginal utility of income is inversely proportional to income.

In the following sections, we explore the importance of allowing for richer forms of income effects. We first discuss implications for consumer welfare and for the curvature of consumer demand. Then we develop an empirical application to the market for butter and margarine and show that income effects are important for estimating accurately how individual demand elasticities and welfare effects depend on y.

3.2.1 Welfare

One important use of random utility models is to compute the welfare impacts of a change in prices, product characteristics or choice sets. In industrial organization, the focus often is on the impact on welfare of price changes (for example, due to a merger as in Nevo (2000), or due to the introduction of a tax as in Kim and Cotterill (2008)). In environmental economics, the focus is on the impact of a change in environmental amenities. In transport economics, the focus is on public investments in transport infrastructure or on taxes or subsidies that affect various modes of transport.

In the vast majority of applications of discrete choice demand models that explicitly compute consumer welfare changes researchers use the linear utility specification (as specified in equation (3.4)) including income in the model as a preference shifter (as in equation (3.5)).[1] In this case, measuring consumer welfare changes is relatively straightforward. In particular, the change in consumer welfare associated with a policy change is invariant to whether it is evaluated before or after the logit shocks, $\boldsymbol{\varepsilon}$, are realized, and can be computed (conditional on realizations of any random coefficients) using the formula derived by Small and Rosen (1981).

When utility is specified as a nonlinear function of $y - p_j$, consumer welfare depends on whether it is evaluated prior to or after the logit shocks are realized (McFadden (1999)). If the logit shocks represent genuine uncertainty from the consumer perspective, it may be appropriate

[1] Petrin (2002) is an exception. He use the log specification given by equation (3.6), and he estimates the consumer welfare effects of the introduction of minivans to the automobile market.

to use an ex ante welfare criterion based on the individual consumer's *expected utility* prior to observing $\boldsymbol{\varepsilon}$. In this case, aggregate welfare is the sum of the individual expected utilities. Conversely, if there is no uncertainty for the consumer over $\boldsymbol{\varepsilon}$ but rather the logit shocks simply represent cross-sectional unobserved heterogeneity, then consumer welfare changes should be based on an ex post criterion based on the individual consumer's *realized utility.* In this case, aggregate welfare is the sum or average of the individual's realized utilities. We present results adopting the latter perspective, based on realized utilities. Like Herriges and Kling (1999) we find in our application that both views yield similar estimates.

Consider baseline prices $\mathbf{p}$ and counterfactual prices $\mathbf{p}'$ (for instance, associated with the introduction of a tax). We measure the change in consumer welfare using compensating variation - the monetary amount required to compensate the consumer post policy change that would make them indifferent to the change.[2] Individual level compensating variation, cv, associated with the price change satisfies

$$V\left(\mathbf{p}, y, \mathbf{x}, \boldsymbol{\varepsilon}\right) = V\left(\mathbf{p}', y - cv, \mathbf{x}, \boldsymbol{\varepsilon}\right). \tag{3.7}$$

Individual cv depends on $\boldsymbol{\varepsilon}$ and therefore is a random variable from the point of view of the econometrician. From the econometrician's perspective, aggregate welfare is the average value of cv, $CV = \mathbb{E}(cv)$.

McFadden (1999) and Herriges and Kling (1999) develop Monte Carlo Markov chain simulation methods that allow for computation of CV in the case of a nested logit model with income effects. More recently Dagsvik and Karlström (2005) have exploited duality results applied to random utility models to characterize the distribution of cv for general random utility models. Using their methods, computation of compensating variation reduces to repeated computation of a one dimensional integral. We use their results to eliminate simulation error in computing CV at the cost of much higher computational effort.

3.3 Market level demand and pass-through

A number of papers have highlighted that the curvature of market demand is a crucial determinant of pass-through of cost shocks and taxes (see, *inter alia*, Seade (1985), Anderson et al. (2001) and Weyl and Fabinger (2013)). Weyl and Fabinger (2013) emphasize that, in the context of a monopolist or symmetrically differentiated single product firm oligopoly, the curvature of the log of demand is key. A simple example illustrates the point.

Consider a single product monopolist with constant marginal cost, c. Let the demand curve be $q(p)$. Optimization implies $q+p\frac{dq}{dp} = c\frac{dq}{dp}$. Differentiating with respect to cost and substituting

[2] The analysis is similar for equivalent variation. When conditional utility is nonlinear in $y-p_j$, the numerical values of compensating and equivalent variation will differ.

yields pass-through as

$$\frac{dp}{dc} = \frac{1}{2 - q\frac{d^2q}{dp^2}/(\frac{dq}{dp})^2} = \frac{1}{1 - \left(\frac{d^2 \ln q}{dp^2}\right)\left(q/\frac{dq}{dp}\right)^2}.$$

This expression shows that pass-through will be incomplete ($\frac{dp}{dc} < 1$) if and only if demand is log-concave ($\frac{d^2 \ln q}{dp^2} < 0$). In this case, restricting market demand to be log-concave rules out pass-through exceeding 100% by assumption. More generally, assuming a particular degree of concavity or convexity of log demand will not necessarily imply under or over-shifting exactly, but may nonetheless place strong restrictions on the possible range of pass-through. In particular, in the logit demand model, heterogeneity in consumer types and the functional form of $\widetilde{U}(y-p_j, \mathbf{x_j})$ both have a strong bearing on the permissible curvature of the log of market demand, and therefore on pass-through.

We develop these ideas in the context of the market demand curve allowing for individual heterogeneity in both income and preferences. Let each consumer be indexed by (y, θ) where as discussed above y measures income and θ measures all other observable and unobservable consumer attributes that enter into utility. Normalizing the size of the market to be one, the market demand curve for option j is then given by:

$$q_j(\mathbf{p}) = \int P_j(y, \theta) g(y, \theta) dy d\theta, \tag{3.8}$$

where $P_j(y, \theta)$ is the individual purchase probability (in the logit case this is given by equation (3.3)) and $g(y, \theta)$ is the joint density over the elements of (y, θ). The second derivative of the log of market demand with respect to price is given by:

$$\begin{aligned}\frac{\partial^2 \ln q_j}{\partial p_j^2} = &\int \frac{P_j(y,\theta)}{q_j} \frac{\partial^2 \ln P_j(y,\theta)}{\partial p_j^2} g(y,\theta) dy d\theta \\ &+ \left[\int \frac{P_j(y,\theta)}{q_j}\left(\frac{\partial \ln P_j(y,\theta)}{\partial p_j}\right)^2 g(y,\theta) dy d\theta - \left(\int \frac{P_j(y,\theta)}{q_j}\frac{\partial \ln P_j(y,\theta)}{\partial p_j} g(y,\theta) dy d\theta\right)^2\right].\end{aligned} \tag{3.9}$$

The curvature of the log of market demand depends on two terms. The first term is the probability weighted average of the second derivatives of log individual demand. The second term is the probability weighted variance of the slope of log individual level demand. The first term is negative if individual level demand is log-concave. The second term is non-negative and is positive when there is heterogeneity in individual demands. Log demand will be concave if individual demand is log-concave and if the cross-sectional variance of the slope of log demand is not too big. It will be convex if individual log demand is convex or if the variance term is large enough in magnitude.

In the case of a linear utility logit model with is no heterogeneity, $\frac{\partial^2 \ln q_j}{\partial p_j^2}$ collapses to the second derivative of the log of individual level demand:

$$\frac{\partial^2 \ln q_j}{\partial p_j^2} = \frac{\partial^2 \ln P_j}{\partial p_j^2} = -\alpha^2 P_j(1-P_j) < 0.$$

The curvature of the log of market demand is then fully determined by the marginal utility of income and the market share. Both individual and market demand are restricted to be log-concave. Adding heterogeneity in consumer preferences maintains the restriction on individual demand but allows for the possibility that the market demand curve might be log-convex or even be log-concave in some regions and log-convex in others.

Allowing $y - p_j$ to enter utility in a flexible nonlinear way relaxes restrictions on the curvature of both individual level and market demand. In particular, with nonlinear utility, individual level demand need not be constrained to be log-concave. The second derivative of the log of consumer demand for option j with respect to its own price is given by:

$$\frac{\partial^2 \ln P_j}{\partial p_j^2} = (1-P_j)\left[\frac{\partial^2 \widetilde{U}(y-p_j,\mathbf{x_j})}{\partial (y-p_j)^2} - \left(\frac{\partial \widetilde{U}(y-p_j,\mathbf{x_j})}{\partial (y-p_j)}\right)^2 P_j\right]. \tag{3.10}$$

The degree of log-concavity (or convexity) is determined by the shape of the function $\widetilde{U}$, and therefore the flexibility of the curvature of individual demand depends on the flexibility of the function $\widetilde{U}$. If $y - p_j$ enters utility in logs as in equation (3.6), the curvature of consumer level demand is very restricted, and is log concave. However, more flexible forms of the function $\widetilde{U}$ allow for more flexibility in consumer level demand curvature including the possibility that consumer demand is log-convex in some regions (individual demand will be log-convex if $\widetilde{U}$ is sufficiently convex). Therefore, specifying utility to be a flexible nonlinear function of $y - p_j$ allows for flexibility in the curvature of market demand both through influencing the variance of the slope of individual demands and through relaxing curvature restrictions on individual demands.

3.4 Illustrative application

To illustrate the potential importance of modeling income effects in a flexible way we provide an example using the UK market for butter and margarine. We have purposely chosen a market that represents a small share of expenditure (this market accounts for just over 1% of households' regular grocery expenditure). The expectation is that income effects play a limited role in this market. We estimate demand under a number of different assumptions about the nature of income effects. We compute individual and market level demand elasticities and simulate the impact of an excise tax. We compare the tax pass-through and consumer welfare predictions of the various specifications.

3.4.1 Consumers

Let i index consumers and t denote time. The index $j \in \{1, ..., J\}$ indexes butter and margarine products. We define a product as a brand-pack size combination and index brands by $b = 1, ..., B$. Product $j = 0$ is the outside option. The product characteristics are given by $\mathbf{x_{jt}} = (a_{bt}, z_j, \mathbf{w_b}, \xi_b)$. a_{bt} measures advertising expenditure for brand b in period t. z_j is pack size, $\mathbf{w_b}$ is a vector of observable brand characteristics and ξ_b is an unobserved (by the econometrician) brand characteristic. In our application, a_{bt} varies over time but not within brand, z_j varies within brand but not over time, and (w_b, ξ_b) do not vary over time. We discuss how this variation contributes to identification in Section 3.4.2 below.

We assume preferences for groceries are weakly separable from other goods and measure y_i as consumer i's average weekly grocery expenditure over a calendar year. By grocery expenditure we mean the household's total expenditure on fast-moving consumer goods, these are products bought in supermarkets and taken home (including food, cleaning products and toiletries).

We assume utility from selecting butter or margarine product j takes the form

$$U_{ijt} = f(y_i - p_{jt}; \alpha_i) + \gamma a_{bt} + \lambda_i z_j + \beta_i' \mathbf{w_b} + \xi_b + \varepsilon_{ijt}, \tag{3.11}$$

and utility from selecting the outside option is given by

$$U_{i0t} = f(y_i; \alpha_i) + \varepsilon_{i0t}. \tag{3.12}$$

We specify a number of different forms for $f(y_i - p_{jt}; \alpha_i)$.

- Polynomial utility
$$f(y_i - p_{jt}; \alpha_i) = \alpha_i^{(1)}(y_i - p_{jt}) + \sum_{n=2}^{N} \alpha^{(n)}(y_i - p_{jt})^n$$
- Linear utility
$$f(y_i - p_{jt}; \alpha_i) = \alpha_i^{(1)}(y_i - p_{jt})$$
- Preference shifter
$$f(y_i - p_{jt}; \alpha_i) = (\alpha_i^{(1)} + \alpha^y y_i)p_{jt}$$

- Log utility

$$f(y_i - p_{jt}; \alpha_i) = \alpha_i^{(1)} \ln(y_i - p_{jt})$$

- Spline utility

$$f(y_i - p_{jt}; \alpha_i) = \sum_{k=1}^{K} a_i^{(k)} B^{(k)}(y_i - p_{jt})$$

where $B^{(k)}(y_i - p_{jt})$ are a set of cubic B-splines with $K-2$ knots placed at the extremes and at the equally spaced percentiles of the expenditure distribution

The first and last specifications include $y_i - p_{jt}$ in utility in a flexible nonlinear way, and therefore admit flexible forms of income effects. We focus our analysis on the polynomial utility case to highlight the role of income effects and to compare more easily with specifications commonly used in the literature. We find empirically that the estimated cubic polynomial utility model closely mimics the estimated spline utility model. In our application, the spline utility estimates are similar to the cubic utility estimates but are more costly to compute.

The linear utility specification rules out income effects; it assumes that the marginal utility of income is constant. It also does not allow purchase patterns to be correlated with total expenditure. The preference shifter specification also assumes the marginal utility of income is constant but does allow the price parameter to shift linearly with total expenditure across households. The log utility model specifies utility to be nonlinear in $y_i - p_{jt}$ and so admits income effects, but $y_i - p_{jt}$ enters utility much less flexibly than in the polynomial or spline utility specifications. We show below that this specification performs very poorly in our application.

The coefficients on the first order $y_i - p_{jt}$ term, pack size and observable brand attributes are allowed to vary across consumers. In particular, we model these coefficients as:

$$\begin{aligned}
\alpha_i^{(1)} =& \alpha_0^{(1)} + \nu_i^{(\alpha)} \\
\lambda_i =& \lambda_0 + \lambda_1 d_i \\
\beta_i =& \beta_0 + \beta_1 d_i + \nu_i^{(\beta)}
\end{aligned}$$

where d_i represents observable consumer demographics. We assume $\nu = (\nu_i^{(\alpha)}, \nu_i^{(\beta)})' \sim N(\mathbf{0}, \Sigma)$ and is uncorrelated with y_i and d_i. We control for the unobserved brand attribute by including brand fixed effects. β_0 is therefore absorbed into the brand fixed effects. Unobserved preference heterogeneity incorporated through the random coefficients allows for correlations in the unobserved portion of consumer utility across options and choice occasions and is crucial in enabling logit choice models to capture realistic substitution patterns across options (see, for instance, Train (2003) and Berry et al. (1995)).

3.4.2 Identification

A common concern in empirical demand analysis is that the ceteris paribus impact of price on demand may not be identified because there may be unmeasured demand shocks that are correlated with price. The most common concern is that price might be correlated with an unobserved product effect, either some innate unobserved characteristic of the product or some market specific shock to demand for the product. Failure to control for the unobserved product effect will lead to inconsistent estimates of the true price effect.

Due to special features of the UK market for butter and margarine and due to the richness of our product level data, we believe the identification strategy proposed in Bajari and Benkard (2005) is reasonable. In particular, we exploit two forms of price variation to identify the marginal utility of income. Firstly, conditional on brand fixed effects and advertising, we argue that it is reasonable to assume that there is no variation in ξ_b (i.e. the relative desirability of one brand over a second does not fluctuate throughout the year). We also argue that the fact that the UK retail food market is characterized by close to national pricing limits the risk that the individual specific demand innovations, ε_{ijt}, conditional on advertising, are correlated with prices.[3] We therefore exploit time series variation in the set of prices. Secondly, we exploit differential nonlinear pricing within brands. This second source of price variation is important in our case and is not exploited in most applications in the literature. Unlike most applications our very detailed data allows us to model demand at the product, rather than brand, level. We are therefore able to include a full set of brand and pack size effects, while exploiting cross-sectional price variation due to differential nonlinear pricing within brands

In the specifications in which utility is a nonlinear function of net expenditure, $y_i - p_{jt}$, we are able to exploit an additional source of variation, the large cross-sectional variation in grocery expenditure across households, to identify the marginal utility of income. This large cross-sectional variation allows us to estimate a very flexible model of income effects. To avoid endogeneity concerns about trip-level grocery expenditure, we measure household expenditure as the household's average weekly expenditure over the course of a calendar year. If we were to measure grocery expenditure at the shopping trip level, a concern might be that trip level expenditure is correlated with idiosyncratic errors in butter and margarine demand (a demand shock leading to the purchase of a particularly expensive butter product would be correlated, all else equal, with higher trip grocery expenditure). A second issue might be that much of the

[3]In the UK most supermarkets implement a national pricing policy following the Competition Commission's investigation into supermarket behavior (Competition Commission (2000)).

high frequency variation in trip level expenditure might reflect planning decisions related to how many shopping trips to undertake in a given period of time and would not be informative of income effects. Use of average weekly expenditure minimizes these concerns and ensures that we only exploit variation in total expenditures which reflects long run expenditure decisions. This will be valid if unobserved preferences that affect substitution within the butter and margarine market are independent from factors that affect average weekly expenditure on all groceries.

3.4.3 Firm competition

Let $f = \{1, ..., F\}$ index firms and F_f denote the set of products owned by firm f. We assume that firms compete by simultaneously setting prices in a Nash-Bertrand game. We consider a mature market with a relatively stable set of products, and we therefore abstract from entry and exit of firms and products from the market. We deploy the commonly used approach of using our demand estimates and an equilibrium pricing condition to infer firms' marginal costs (see Berry (1994) or Nevo (2001)).

Normalizing the size of the market to be one, firm f's (variable) profits in market t are given by:

$$\Pi_{ft}(\mathbf{p_t}) = \sum_{j \in F_f} (p_{jt} - c_{jt}) q_j(\mathbf{p_t}). \tag{3.13}$$

The first order conditions for firm f are

$$q_j(\mathbf{p_t}) + \sum_{k \in F_f} (p_{kt} - c_{kt}) \frac{\partial q_k(\mathbf{p_t})}{\partial p_{jt}} = 0 \quad \forall j \in F_f. \tag{3.14}$$

In a Nash equilibrium, the first order conditions (3.14) are satisfied for all firms. Under the assumption that observed market prices are an equilibrium outcome of the Nash-Bertrand game played by firms, given our estimates of the demand function, we can invert firms' first order conditions to infer marginal costs.

3.4.4 Counterfactual

We simulate the introduction of an excise tax (t) that is proportional to the saturated fat content of a product. Let η_j denote the saturated fat content of product j and $\boldsymbol{\eta} = (\eta_1, ..., \eta_J)'$. A counterfactual equilibrium price vector $\mathbf{p_t^e}$ satisfies:

$$q_j(\mathbf{p_t^e} + t\boldsymbol{\eta}) + \sum_{k \in F_f} (p^e_{kt} - c_{kt}) \frac{\partial q_k(\mathbf{p_t^e} + t\boldsymbol{\eta})}{\partial p_{jt}} = 0 \quad \forall j \in F_f, \text{ and } \forall f \in 1, ..., F. \tag{3.15}$$

In Appendix Section B.1 we show that our analysis and results yield similar conclusions if instead we consider an ad valorem tax (τ_{av}), such that a counterfactual equilibrium price vector

$\mathbf{p_t^e}$ satisfies:

$$q_j((1+\tau_{av}\boldsymbol{\eta})\mathbf{p_t^{av}})+\sum_{k\in F_f}(p_{kt}^{av}-c_{kt})\frac{\partial q_k((1+\tau_{av}\boldsymbol{\eta})\mathbf{p_t^{av}})}{\partial p_{jt}}=0 \quad \forall j\in F_f,\ \text{and}\ \forall f\in 1,...,F. \quad (3.16)$$

3.4.5 Data

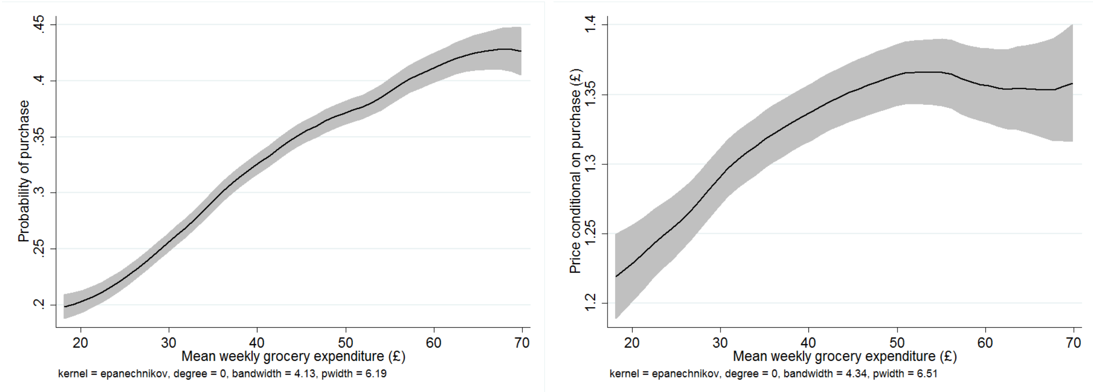
We apply the model to the UK market for butter and margarine. We use purchase date on 10,012 households from Kantar WorldPanel for calendar year 2010. For each household we observe all grocery products that are bought and taken into the home. We define a 'choice occasion' as a household's weekly grocery purchases. We use information on five randomly chosen choice occasions for each household (50,060 in total) to estimate the model. Households purchase a butter or margarine product on 34% of choice occasions. On these choice occasions households on average spend around £1.35 (or 3.5% of their grocery expenditure) on butter and margarine. On the remaining occasions they select the outside option of not purchasing butter or margarine.

We measure grocery expenditure, y_i, as a household's mean weekly grocery expenditure over 2010. As discussed in Section 3.4.2 using average weekly expenditure ensures that our expenditure measure is not correlated with idiosyncratic trip-specific demand shocks. Average expenditure in the sample is £40. The 10th percentile of the distribution is £20 and the 90th percentile is £63.

We provide some reduced form evidence that, despite the fact that butter and margarine are small budget share items, household butter and margarine purchase behavior is correlated with their mean weekly grocery expenditure. The left side panel of Figure 3.1 shows results from a non-parametric regression of the average probability that a household purchases any butter or margarine product on a choice occasion on its mean weekly grocery expenditure. The right side panel displays results from a non-parametric regression of the average price a household pays for butter or margarine, conditional on purchasing, on its mean weekly grocery expenditure. The figures show that higher mean weekly grocery expenditure is strongly correlated with both the probability of purchase, and conditional on purchase, the price of the product chosen. Table 3.1 shows that this pattern remains after conditioning on household size. The pattern is not simply a reflection of correlations between household size, expenditure, and purchase patterns.

Figure 3.1: *Correlation of purchase patterns with mean weekly grocery expenditure:*

A) probability of purchase | *B) price paid conditional on purchase*



Notes: The figures display results from weighted kernel regressions across 10,012 households. The weights ensure the sample is representative of the British population. The left panel shows results from a regression of households' mean probability of purchasing butter or margarine on mean weekly grocery expenditure. The right panel shows results from a regression of households' mean price paid for butter or margarine conditional on purchase on mean weekly grocery expenditure. The shaded areas depict pointwise 95% confidence intervals.

Table 3.1: *Variation in purchase behavior with mean weekly grocery expenditure by household size*

Quartile of expenditure distribution	Probability of purchasing butter margarine	Price paid conditional on purchase
One person households		
1	0.19	0.22
2	0.23	0.29
3	0.23	0.32
4	0.30	0.43
Two person households		
1	0.26	0.32
2	0.35	0.46
3	0.39	0.56
4	0.42	0.60
Three person households		
1	0.27	0.33
2	0.35	0.48
3	0.38	0.52
4	0.45	0.65
Four person households		
1	0.28	0.34
2	0.36	0.47
3	0.42	0.54
4	0.45	0.63

Notes: Quartiles are defined for the distribution of mean weekly grocery expenditure within each household size category.

The butter and margarine market in the UK has an oligopolistic structure. There are eight main firms in the market. Unilever is the largest, marketing 17 products that together have a market share of 52%. The second largest is Dairy Crest with a market share of 26%, followed by Arla with 17%, and Tesco with 3%.

The first four columns of Table 3.2 list the firms that operate in the market, the brands that these firms sell, the pack sizes that each brand is available in and the products the firms sell. In most cases, a product (i.e. an option in a consumer's choice set) is defined as a brand-pack size combination.[4] Column five shows the quantity share of each product. Column six shows the mean market price computed as the transaction weighted mean price in each month. The remaining columns show how the product characteristics we include in the model vary across products. Characteristics include whether the product is butter or margarine, the amount of saturated fat per 100g and monthly advertising expenditure for the brand (in addition to pack size and brand effects).

[4]In a few instances a brand-pack size contains two products - a salted and unsalted version.

Table 3.2: *Products and characteristics*

Firm	Brand	Pack size (Kg)	Product	Quantity share (%)	Price (£)	Butter or margarine	Saturated fat per 100g (g)	Advertising (£m)
Adams	Kerrygold	0.25	Kerrygold 250g	1.00	1.09	Butter	48.93	0.22
Arla				17.76				
	Anchor	0.25	Ar: Anchor 250g	1.48	1.12	Butter	54.00	0.81
	Anchor spr.	0.25	Ar: Anchor spr. 250g	0.38	1.40	Butter	31.20	0.10
	Anchor spr.	0.50	Ar: Anchor spr. 500g	3.03	1.99	Butter	31.20	0.10
	Anchor spr. lighter	0.50	Ar: Anchor light spr. 500g	1.20	1.99	Butter	23.70	0.03
	Lurpak	0.25	Ar: Lurpak ss 500g	0.82	1.17	Butter	52.00	0.73
	Lurpak	0.25	Ar: Lurpak us 250g	0.36	1.17	Butter	53.00	0.73
	Lurpak spr.	0.25	Ar: Lurpak spread ss 250g	0.59	1.42	Butter	36.70	0.04
	Lurpak spr.	0.50	Ar: Lurpak spread ss 500g	5.10	2.15	Butter	36.70	0.04
	Lurpak spr. lighter	0.50	Ar: Lurpak light ss 500g	3.48	2.17	Butter	25.80	0.00
	Lurpak spr. lighter	0.25	Ar: Lurpak light ss 250g	0.55	1.42	Butter	25.80	0.00
Asda	Asda	0.25	Asda 250g	0.73	0.92	Butter	54.00	0.00
Dairy Crest				25.39				
	Clover diet low fat spr.	0.50	DC: Clover diet 500g	1.46	1.26	Margarine	23.50	0.00
	Clover spr.	0.50	DC: Clover spr. 500g	5.21	1.20	Margarine	26.90	0.78
	Clover spr.	1.00	DC: Clover spr. 1kg	4.34	2.37	Margarine	26.90	0.78
	Country Life	0.25	DC: Country life 250g	1.16	1.08	Butter	54.00	0.47
	Country Life	0.25	DC: Country life us 250g	0.39	1.07	Butter	54.70	0.47
	Country Life light spr.	0.50	DC: Country life spr. 500g	1.08	1.87	Butter	23.00	0.00
	Country Life spr.	0.50	DC: Country life spr. 500g	1.24	2.13	Butter	31.40	0.00
	Utterly Butterly	0.50	DC: Utterly Butterly 500g	5.99	0.80	Margarine	14.70	0.15
	Utterly Butterly	1.00	DC: Utterly Butterly 1kg	1.36	1.95	Margarine	14.70	0.15
	Vitalite	0.50	DC: Vitalite 500g	2.17	0.92	Margarine	13.40	0.00
	Willow	0.25	DC: Willow 250g	0.97	0.67	Margarine	17.06	0.00
Sainsburys				0.91				
	Sainsburys	0.25	Sainsburys 250g	0.66	0.93	Butter	54.00	0.00
	Sainsburys	0.25	Sainsburys us 250g	0.24	0.93	Butter	54.00	0.00
Morrisons	Morrisons	0.25	Morrisons 250g	0.36	0.90	Butter	52.10	0.00
Tesco				3.15				
	Tesco butter me up	0.50	Tesco butter me up 500g	1.06	0.97	Margarine	17.50	0.00
	Tesco blended	0.25	Tesco blended 250g	0.34	1.03	Butter	48.60	0.00
	Tesco value	0.25	Tesco value 250g	1.43	0.92	Butter	48.60	0.00
	Tesco value	0.25	Tesco value us 250g	0.32	0.93	Butter	48.60	0.00
Unilever				51.54				
	Bertolli light olive spr.	0.50	Un: Bertolli light 500g	1.49	1.35	Margarine	9.50	0.21
	Bertolli spr.	1.00	Un: Bertolli 1kg	1.74	2.66	Margarine	14.00	1.00
	Bertolli spr.	0.50	Un: Bertolli 500g	2.59	1.34	Margarine	14.00	1.00
	Flora buttery	0.50	Un: Flora buttery 500g	7.53	1.02	Margarine	15.60	0.76
	Flora extra light spr.	0.50	Un: Flora extra light 500g	0.98	1.34	Margarine	5.10	0.60
	Flora light spr.	0.50	Un: Flora light 500g	3.55	1.22	Margarine	9.30	0.60
	Flora light spr.	1.00	Un: Flora light 1kg	4.64	2.29	Margarine	9.30	0.60
	Flora ProActiv spr.	0.25	Un: Flora proactive 250g	0.46	1.87	Margarine	8.00	1.37
	Flora ProActiv spr.	0.50	Un: Flora proactive 500g	0.82	3.62	Margarine	8.00	1.37
	Flora	0.50	Un: Flora 500g	2.37	1.22	Margarine	12.00	0.11
	Flora	1.00	Un: Flora 1kg	2.45	2.30	Margarine	12.00	0.11
	ICBINB	1.00	Un: ICBINB 1kg	2.02	2.02	Margarine	19.90	0.35
	ICBINB	0.50	Un: ICBINB 500g	7.72	0.85	Margarine	19.90	0.35
	ICBINB light	0.50	Un: ICBINB light 500g	3.70	0.87	Margarine	11.00	0.11
	Stork baking block	0.25	Un: Stork 250g	0.66	0.50	Margarine	25.70	0.19
	Stork pack	0.50	Un: Stork 500g	3.00	0.72	Margarine	14.80	0.11
	Stork pack	1.00	Un: Stork 1kg	5.82	1.34	Margarine	14.80	0.11

Notes: Price and advertising are unweighted means across all 12 markets. Quantity shares are computed using 50,060 observations in our data, weighted to reflect the British population.

3.4.6 Estimates

We estimate the five specifications - polynomial utility, linear utility, preference shifter, log utility, and spline utility - outlined in Section 3.4.1 using maximum likelihood.[5] In Table 3.3 we report the coefficient estimates.[6] For the polynomial utility specification, we specify utility as a third order polynomial of net expenditure, $y - p_j$. As shown in Figure 3.2 this provides sufficient flexibility to capture the shape of the conditional marginal utility of income implied by the estimates of the spline utility specification.

The top panel presents estimates of the random coefficients. For each specification we model the coefficient on the first order net expenditure term as a random coefficient, meaning we allow for unobserved preference heterogeneity across households. We also include a random coefficient on the attribute "butter". As "butter" is collinear with the brand effects, we constrain it to have zero mean, but we allow the mean to shift with whether the main shopper has a body mass index indicating he or she is obese. We assume that the random coefficients are joint normally distributed and allow for correlation between the coefficients. Direct interpretation of the $y - p_j$ coefficients is difficult. We simply note that the means of all $y - p_j$ coefficients are statistically significant, as are all the higher order, interaction, variance and covariance parameters.

The bottom section of the table shows the coefficient estimates for the non-random coefficients. In each case the advertising coefficient is statistically insignificant indicating little evidence that butter and margarine advertising has a strong contemporaneous impact on demand. We interact pack size effects with household size, which captures the fact that larger households are more likely to select large pack sizes. We also interact pack size with a dummy indicating whether the main shopper is obese. Three of the four specifications indicate obese main shoppers have a statistically significant preference for 500g and 1kg pack sizes over the smaller 250g pack size (the coefficients are not statistically significant in the log utility model). Like the butter dummy, a product's saturated fat content per 100g is collinear with the brand effects. Therefore we include this attribute interacted with the obese dummy. In all four specifications the obese-butter interaction is positive and statistically significant and the obese-saturated fat interaction is not statistically significant, indicating obese consumers have a stronger preference than other consumers for butter, but conditional on their preference for butter, they do not prefer products with higher saturated fat.

[5] We use Gauss-Hermite quadrature rules to eliminate simulation error when computing the likelihood function.

[6] For brevity, we do not show spline results in Table 3.3. Results from the spline utility model are statistically indistinguishable from the cubic utility model. In Figure 3.2 we illustrate this point by graphing the estimated conditional marginal utility of income from both the spline utility and cubic utility models.

Table 3.3: *Coefficient estimates*

	Polynomial utility		Linear utility		Preference shifter		Log utility	
	Coefficient estimate	Standard error	Coefficient estimate	Standard error	Coefficient estimate	Standard error	Coefficient estimate	Standard error
Random coefficients								
Mean terms								
(y-p)	3.8134	0.0888	2.5238	0.0507	3.1991	0.0602		
ln(y-p)							4.4524	0.1443
Higher order terms								
$(y\text{-}p)^2$	-0.2352	0.0164						
$(y\text{-}p)^3$	0.0115	0.0012						
Interaction terms								
(y-p)*y					-0.1512	0.0067		
Butter*Obese	0.1301	0.0713	0.1482	0.0720	0.1406	0.0712	0.1482	0.0975
Variance-covariance terms								
Var(y-p)	0.7410	0.0149	0.7851	0.0151	0.7515	0.0148		
Var(ln(y-p))							6.0865	0.3788
Var(Butter)	1.7198	0.0333	1.8075	0.0324	1.7332	0.0331	4.5884	0.2307
Cov(y-p, Butter)	1.4204	0.0466	1.3610	0.0448	1.4083	0.0461		
Cov(ln(y-p), Butter)							2.9557	0.2026
Fixed coefficients								
Advertising	-0.0307	0.0224	-0.0305	0.0224	-0.0306	0.0224	-0.0180	0.0321
500g	3.0214	0.0640	2.7368	0.0634	3.0053	0.0640	2.0281	0.0692
1kg	3.6145	0.1190	3.1601	0.1178	3.6073	0.1191	1.2973	0.1187
500g*HHsize	0.0796	0.0109	0.1887	0.0102	0.0840	0.0109	0.0109	0.0134
1kg*HHsize	0.2629	0.0221	0.4343	0.0211	0.2659	0.0221	0.1974	0.0266
500g*Obese	0.1321	0.0298	0.1730	0.0303	0.1409	0.0299	0.0307	0.0381
1kg*Obese	0.1630	0.0592	0.2255	0.0598	0.1760	0.0596	0.0314	0.0740
Satfat*Obese	0.0397	0.1178	0.0582	0.1181	0.0465	0.1178	0.2721	0.1619

Notes: Sample size is 50,060 choice occasions involving 10,012 different households. Random coefficients are assumed to be distributed joint normally. The butter dummy is collinear with the brand effects and therefore has a mean coefficient that is constrained to be zero.

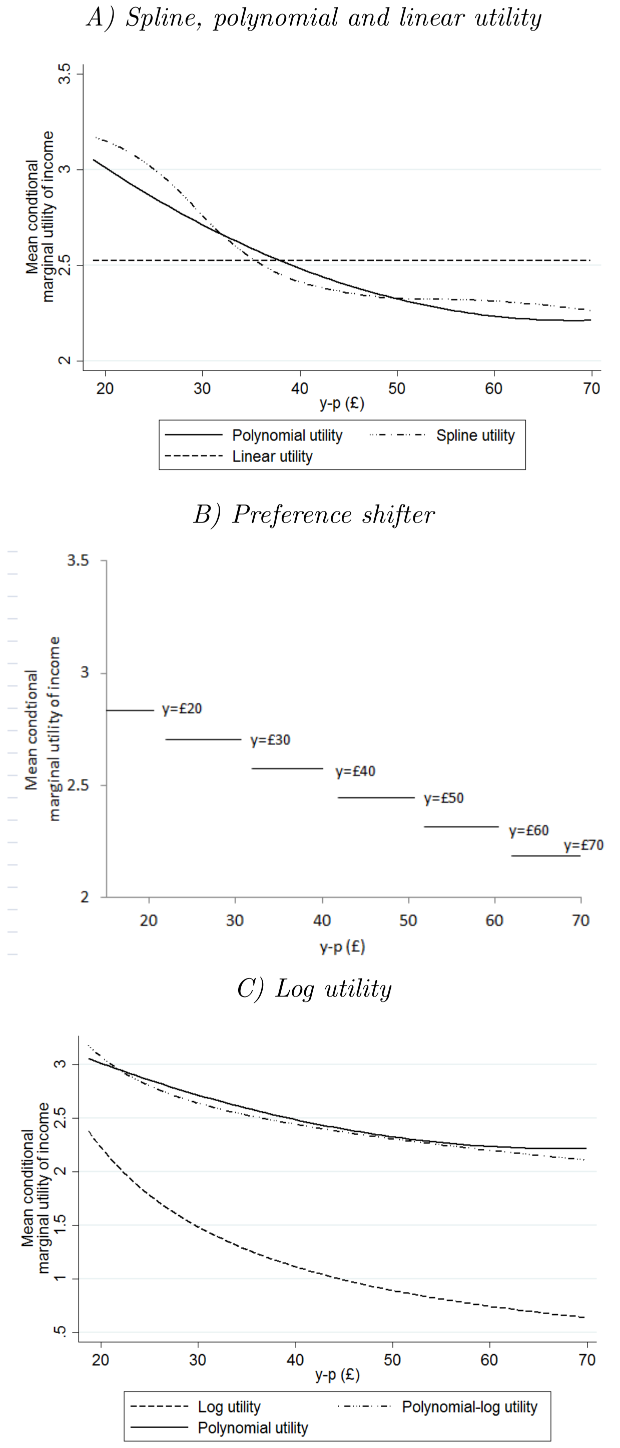
As discussed in Section 3.2, the behavior of the marginal impact of a change in net expenditure, $y - p_j$, on utility is a crucial determinant of both welfare effects and pass-through of tax and cost shocks. In Figure 3.2 we show how the mean conditional marginal utility of income varies with $y - p_j$. Panel A shows estimates for the polynomial utility, spline utility and linear utility specifications, panel B shows numbers for the preference shifter specification and panel C focuses on the log utility specification.

The estimates of the spline utility specification show that the conditional marginal utility of income is a decreasing function of net expenditure and that over most of the domain (excluding the very bottom and top of the expenditure distribution), the function is convex. The cubic polynomial utility specification captures this shape well. In Section 3.3 we highlighted that allowing utility to depend on $y - p_j$ through a nonlinear function, $\widetilde{U}(.)$, allows for the possibility of household level demands that are log-convex (something that is typically ruled out in applied applications). Log-convex demand arises if $\widetilde{U}(.)$ is sufficiently convex, which requires the conditional marginal utility of income to be an increasing function of $y - p_j$. Figure 3.2 makes clear that in our application we do not find evidence of log-convex household demands. The linear utility model constrains the marginal utility of income to be constant and uncorrelated with consumer expenditure. This restriction is clearly not supported by the data.

Like the linear utility specification, the preference shifter specification imposes that the conditional marginal utility of income is constant for a given household. However, it does allow the parameter to shift linearly across households based on their total expenditure, y. Panel B of Figure 3.2 shows that the specification does, to some extent, capture the fact that households with higher total expenditure have a lower conditional marginal utility of income. However, the linear way in which y interacts with the coefficient on price, means the specification is unable to capture the convexity exhibited in the estimates of the polynomial specification.

The log utility specification, shown on panel C of Figure 3.2, yields an estimate of the conditional marginal utility of income that decreases convexly, but the function is shifted vertically downwards compared to the function implied by the polynomial specification (also shown on the graph). In principle this could reflect mis-specification of both the spline and polynomial utility models, or mis-specification of the log utility specification. The latter is much more likely, because specifying utility to be linear in the log of $y - p_j$ leaves only one parameter to determine the location, slope and curvature of the conditional marginal utility of income function. To test whether this is indeed the case we re-estimated the model specifying utility as a *third order polynomial in the log of* $y - p_j$ (denoted polynomial-log utility in the figure). This model, which is more general and nests the log utility specification, yields an estimate of the conditional marginal utility of income that is very similar to the polynomial utility specification.

Figure 3.2: *Conditional marginal utility of income*



Notes: Lines shows mean conditional marginal utility of income after integrating out the random coefficients.

Our counterfactual analysis requires solution of a series of nonlinear first order conditions. Using estimates from the spline utility model results in relatively slow, and in some cases unstable, computations. As our baseline model we therefore proceed with the polynomial specification. It is clear from panel C of Figure 3.2 that, in our application, the log utility specification does a very poor job of replicating the shape of the conditional marginal utility of income found with more flexible specifications. In addition, the log utility model yields implausible estimates of marginal costs and welfare. In what follows we therefore compare our baseline model to the linear utility and preference shifter specifications.

The market level price elasticities are crucial determinants of equilibrium prices in models of firm pricing in imperfectly competitive markets. It turns out in our empirical application that the polynomial utility, linear utility and preference shifter models all yield market level price elasticity and marginal cost estimates that are very similar.[7] In other words, all three specifications agree on the slope of market demand at observed prices. This need not be true in general.

While market elasticities determine the nature of the pricing equilibrium, household level elasticities are important for determining the distributional impact of a policy reform. We find in our application that, unlike the market elasticities, the household level elasticities are sensitive to whether we model income effects in a flexible and theoretically consistent way or not. To illustrate this, we compute each household's own-price elasticity of demand for butter and margarine for each choice occasion in our data (this is the market share weighted average of household's own-price elasticities across products).

In Table 3.4 we report the mean household level own price elasticity under each specification, and we report the average deviation from the mean own price elasticity for households in each quartile of the total expenditure distribution. The table also contains 95% confidence intervals.[8] In Figure 3.3 we plot how household level own price elasticities vary with total expenditure for each of the model specifications. The mean household own price elasticity is essentially the same under each model specification, however the three specifications yield different predictions for how price sensitivity varies across the expenditure distribution. The polynomial utility specification results indicate that households with low expenditure are the most price sensitive; households in the bottom quartile of the expenditure distribution, on average, have an own price elasticity 0.27 below the mean and households in the top quartile, on average, have an own price elasticity 0.21 above the mean. The linear utility model completely fails to capture the variation in price sensitivity across the expenditure distribution, which is not surprising since expenditure plays no role in determining patterns of demand in this specification. The preference shifter specification does predict falling price sensitivity across the expenditure distribution, but it fails

[7] See Appendix Section B.2.

[8] We calculate confidence intervals in the following way. We obtain the variance-covariance matrix for the parameter vector estimates using standard asymptotic results. We then take 100 draws of the parameter vector from the joint normal asymptotic distribution of the parameters and, for each draw, compute the statistic of interest, using the resulting distribution across draws to compute Monte Carlo confidence intervals (which need not be symmetric around the statistic estimates).

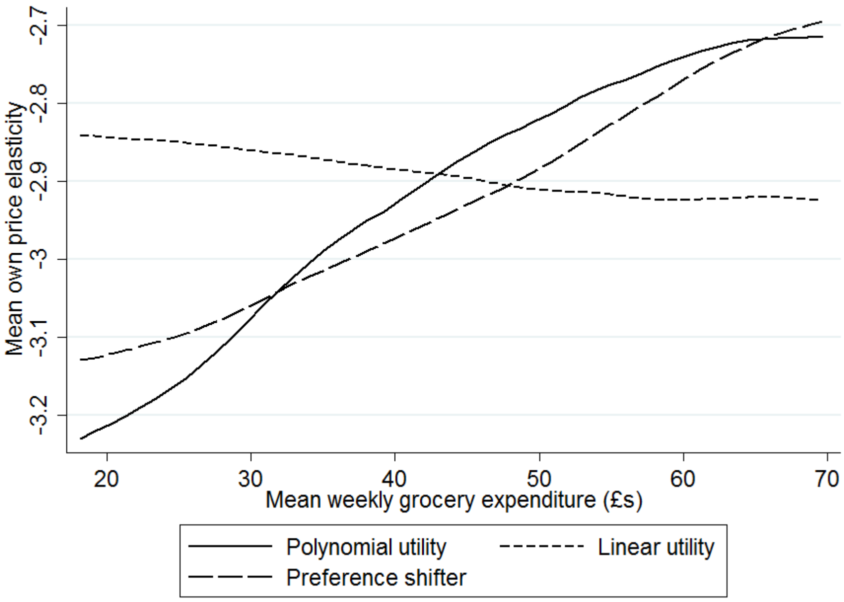
to capture the concavity in the relationship, underestimating price sensitivity at the bottom of the expenditure distribution and overestimating it in the center.

Table 3.4: *Household own price elasticity*

Specification	Mean own price elasticity	Average deviation from mean own price elasticity for quartile of expenditure distribution: 1	2	3	4
Polynomial utility	-2.94	-0.27	-0.04	0.10	0.21
	[-3.04, -2.79]	[-0.30, -0.25]	[-0.05, -0.03]	[0.09, 0.11]	[0.18, 0.23]
Linear utility	-2.89	0.04	0.01	-0.02	-0.04
	[-2.97, -2.75]	[0.03, 0.04]	[0.02, 0.02]	[-0.02, -0.01]	[-0.04, -0.03]
Preference shifter	-2.93	-0.19	-0.07	0.03	0.23
	[-3.06, -2.79]	[-0.22, -0.18]	[-0.08, -0.06]	[0.03, 0.04]	[0.21, 0.27]

Notes: For each choice occasion we compute the market-share weighted mean own price elasticity. Numbers shows average of this own price elasticity. We measure expenditure as the households' mean weekly grocery expenditure. 95% confidence intervals are shown in brackets.

Figure 3.3: *Variation in own price elasticities with expenditure*



Notes: For each choice occasion we compute the market-share weighted mean own price elasticity. Figure shows local polynomial regression of how mean choice occasion elasticity varies with households' mean weekly grocery expenditure.

3.4.7 Counterfactual results

To illustrate how assumptions about the marginal utility of income may affect conclusions about the impact on market equilibria and the welfare effects of policy reform we simulate the effect of an excise tax that is proportional to the saturated fat content of the product (see Section 3.4.4). We select the level of the tax that generates a 25% fall in purchases of saturated fat in the case of no firm pricing response (i.e. in the case of 100% pass-through).

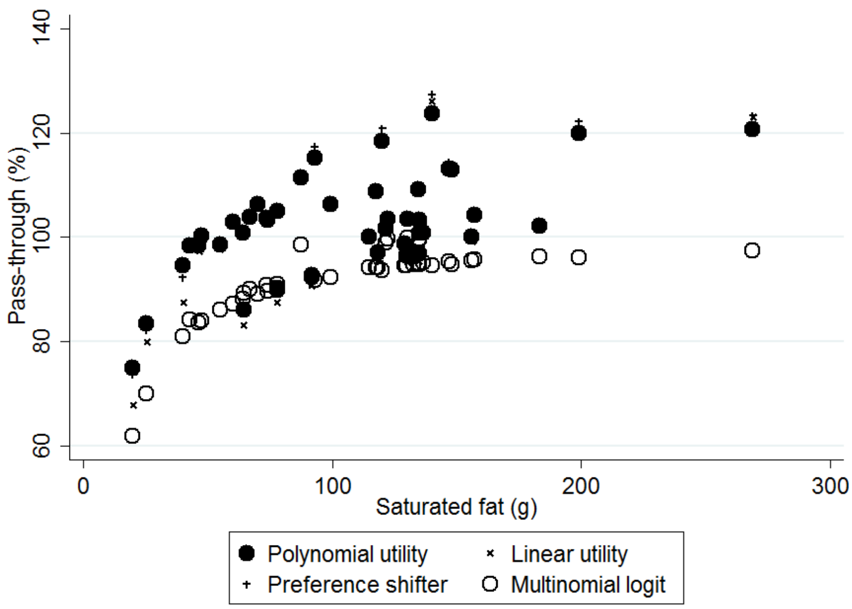
Figure 3.4 is a scatter plot, at the product level, that shows how tax pass-through is related to a product's total saturated fat content. We plot the numbers for the polynomial utility specification and for three alternative specifications - the linear utility and preference shifter specifications and a simple logit specification with linear utility and with no consumer level heterogeneity.

For the polynomial utility specification, across all products in the market, average pass-through of the tax to consumer prices is 103%. Therefore, on average the model predicts that prices will move close to one-to-one with the excise tax. This average masks a considerable degree of heterogeneity across products. Figure 3.4 shows that products with higher saturated fat contents tend to have higher tax pass-through. As the tax is levied on saturated fat content, this implies that firms' equilibrium pricing response acts to amplify the price differential the tax creates between low and high fat products.

In Section 3.3 we highlighted that an important determinant of tax pass-through is the curvature of the log of market demand, and that an advantage of a model in which utility is flexible and nonlinear in $y - p_j$ over commonly used specifications is that it relaxes restrictions on the curvature of log market demand through allowing for more flexibility in the curvature of log household demands. This flexibility allows one to test empirically whether market demand is log-concave or not. Figure 3.4 shows that, in our application, the alternative more restrictive linear utility specification actually yields pass-through results that are very similar to those found by the polynomial utility specification. This is also true for the preference shifter model. In this market, this suggests that the curvature restrictions placed on household level demands (e.g. log-concavity) when utility is linear in $y - p_j$ are not overly restrictive. Together these results provide empirical evidence that market demand is log-concave in this market. If we had only estimated the linear utility model, we would not be able to provide empirical evidence on this question because we would have imposed *a priori* log-concavity.

A second determinant of the curvature of log market demand is the average variance of the slope of the log of household demand curves. In each of the polynomial utility, linear utility and preference shifter specifications we allow for the possibility that the variance is non-zero through the inclusion of unobserved preference heterogeneity (through random coefficients). In addition, the preference shifter and polynomial models also allow for positive variance through the inclusion of expenditure (as a preference shifter in the first case, and as an argument of consumer level utility in the second). Allowing for this heterogeneity is important in practice. Figure 3.4 shows that a multinomial logit model that excludes any preference heterogeneity, and in which utility is specified to be linear in $y - p_j$, yields pass-through which is lower than the random coefficient models; pass-through is 92% on average. It is well know that inclusion of rich preference heterogeneity in logit demand models is important for capturing realistic substitution patterns. Our results suggest, not surprisingly, it is also important when modeling pass-through.

Figure 3.4: *Tax pass-through across products*



Notes: For each product in each market with compute the pass-through of the tax. Figure is a scatter plot of products' mean pass-through across markets with their saturated fat contents.

In the first column of Table 3.5 we report average compensating variation estimated using each model specification. These numbers can be interpreted as the monetary payment (per year) the average household would require to be indifferent to the change in tax policy. All three models predict average compensating variation of around £2.

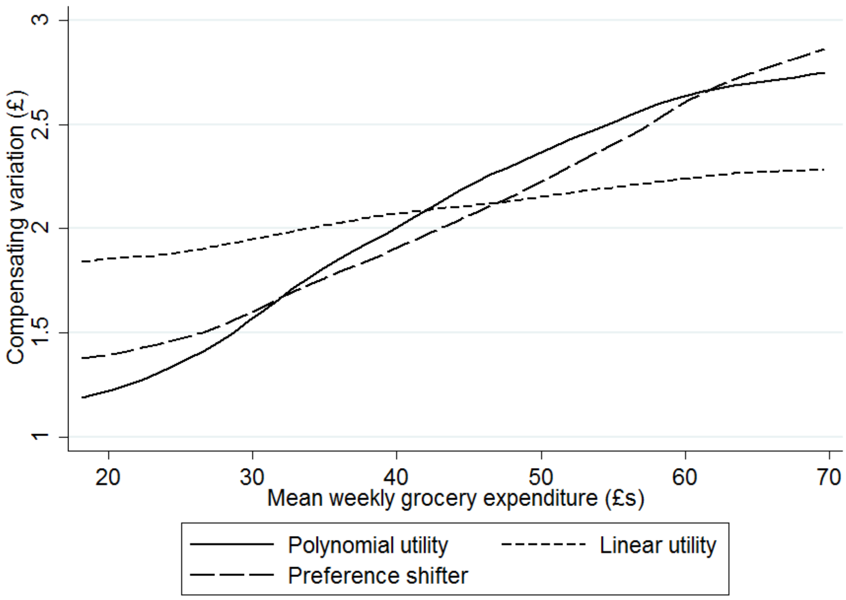
Columns two to five of Table 3.5 show the average deviation from mean compensating variation for households in each quartile of the expenditure distribution. Figure 3.5 shows graphically how compensating variation varies with total expenditure. All model specifications suggest compensating variation is increasing in mean weekly grocery expenditure. For the linear model the increase is comparatively small and is driven by compensating variation being related to household characteristics that are correlated with total expenditure (as the latter drops out of the model). The polynomial utility model suggests that the relationship between compensating variation and total expenditure is much stronger; on average households in the bottom quartile of the expenditure distribution have compensating variation of £0.74 below average and household in the top quartile, on average, have compensating variation £0.67 above average. Households towards the bottom of the expenditure distribution both purchase less butter and margarine and are more willing to switch between alternatives in response to a price change, leading them to be less badly affected in absolute terms than households with higher expenditure. The preference shifter model also predicts a positive relationship between a household's expenditure and compensating variation, however, it fails to capture the concavity of the relationship and overestimates compensating variation at the bottom of the expenditure distribution and underestimates it towards the center.

Table 3.5: *Compensating variation from tax*

Specification	Mean compensating variation	Average deviation from mean compensating for quartile of expenditure distribution: 1	2	3	4
Polynomial utility	2.03	-0.74	-0.19	0.26	0.67
	[1.95, 2.18]	[-0.81, -0.69]	[-0.23, -0.16]	[0.23, 0.30]	[0.60, 0.75]
Linear utility	2.08	-0.24	-0.05	0.04	0.23
	[1.98, 2.22]	[-0.26, -0.21]	[-0.05, -0.04]	[0.04, 0.05]	[0.21, 0.26]
Preference shifter	2.05	-0.63	-0.28	0.09	0.76
	[1.91, 2.20]	[-0.70, -0.56]	[-0.30, -0.23]	[0.10, 0.12]	[0.70, 0.87]

Notes: Numbers give compensating variation for the average household associated with the simulated excise tax. We measure expenditure as the households' mean weekly grocery expenditure. Numbers are for a calendar year. 95% confidence intervals are shown in brackets.

Figure 3.5: *Variation in compensating variation from tax with expenditure*



Notes: Figure shows local polynomial regression of how compensating variation from tax varies with households' mean weekly grocery expenditure.

The distributional results from the preference shifter model differ from those from the polynomial utility specification because the preference shifter model does not allow enough flexibility in the way in which y enters to fully recover how purchase patterns vary with total expenditure.

It is straightforward to demonstrate this empirically, and at the same time suggest a modification to the preference shifter model that allows it to recover the full distributional consequence of the saturated fat tax. If the true model is cubic as our results suggest, a first order approximation around $p_j = 0$ is given by:

$$U_j \approx -(a^{(1)} + a^{(2)}y + a^{(3)}y^2)p_j + g(\mathbf{x_j}) + \epsilon_j \tag{3.17}$$

where we have omitted all terms that do not vary across j and where $a^{(1)} = \alpha^{(1)}$, $a^{(2)} = 2\alpha^{(2)}$ and $a^{(3)} = 3\alpha^{(3)}$. The approximation error is quadratic in p_j and depends on $\widetilde{U}''$. If for a given consumer the conditional marginal utility of income is approximately constant in the region $[y - p_j, y]$ then the approximation will work well. When utility is smooth, this will be the case when p_j is small relative to y. In our application, estimation of the linearized utility model associated with equation (3.17) yields results, including distributional effects, which are very close to those from the cubic polynomial specification. While this model is not as appealing from a theoretical point of view, it may offer a practically expedient way to capture variation across the income distribution. A researcher who did not know the correct functional form for $\widetilde{U}$ could allow the price coefficient to be a nonparametric function of y. Tables 3.6 and 3.7 illustrate for both the household level elasticities and compensating variation.

Table 3.6: *Mean own price elasticity: Polynomial and linearized utility*

Specification	Mean own price elasticity	Average deviation from mean own price elasticity for quartile of expenditure distribution: 1	2	3	4
Polynomial utility	-2.94	-0.27	-0.04	0.10	0.21
	[-3.04, -2.79]	[-0.30, -0.25]	[-0.05, -0.03]	[0.09, 0.11]	[0.18, 0.23]
Linearized utility	-2.94	-0.26	-0.03	0.11	0.20
	[-3.04, -2.79]	[-0.30, -0.25]	[-0.05, -0.02]	[0.09, 0.12]	[0.17, 0.23]

Notes: For each choice occasion we compute the market-share weighted mean own price elasticity. Numbers shows average of this own price elasticity. We measure expenditure as the households' mean weekly grocery expenditure. 95% confidence intervals are shown in brackets.

Table 3.7: *Compensating variation from tax: Polynomial and linearized utility*

Specification	Mean compensating variation	Average deviation from mean compensating for quartile of expenditure distribution: 1	2	3	4
Polynomial utility	2.03	-0.74	-0.19	0.26	0.67
	[1.95, 2.18]	[-0.81, -0.69]	[-0.23, -0.16]	[0.23, 0.30]	[0.60, 0.75]
Linearized utility	2.04	-0.76	-0.21	0.25	0.66
	[1.95, 2.19]	[-0.82, -0.70]	[-0.23, -0.16]	[0.23, 0.31]	[0.60, 0.75]

Notes: Numbers give compensating variation for the average household associated with the simulated excise tax. We measure expenditure as the households' mean weekly grocery expenditure. Numbers are for a calendar year. 95% confidence intervals are shown in brackets.

In Appendix Section B.1 we shows that if we alternatively consider an ad valorem tax, of the form described in Section 3.4.4, we find that this tax is under-shifted. Our conclusions regarding the effects of not modelling income effects flexibly and in a theoretically rigorous way remain very similar.

3.5 Conclusion

In this paper we have explored the importance of relaxing restrictions commonly placed on the marginal utility of income in logit demand models. By far the two most common approaches are either to assume that the marginal utility of income is constant for a given consumer, but to allow it to vary cross-sectionally with demographics including consumer income, or to model income effects by assuming utility is linear in the log of all spending outside the market currently under focus. Both of these approaches heavily constrain income effects (ruling them out in the first case) and unduly restrict demand curvature. Imposing these restrictions prevent the data from providing evidence as to the true shape of the demand curve.

Specifying consumer level utility in the form $U_j = f(y - p_j) + g(\mathbf{x_j}) + \epsilon_j$ for some flexible nonlinear (e.g. polynomial) function $f(.)$ offers three advantages. Firstly, it allows the model to capture any income effects induced by the policy counterfactual under consideration. Secondly, it allows for more flexibility in the curvature of consumer level demands. Thirdly it allows for a richer relationship between expenditure, demand, and welfare.

To explore the empirical importance of these restrictions we consider an application to the UK butter and margarine market. This product category comprises a small fraction of households' budgets and is a category for which flexible modeling of the marginal utility of income may *a priori* not seem to be of first order importance. Yet we show that results from a flexible model differ from results form standard models in important ways. In the case of the log utility specification, it is clear that the shape imposed on the conditional marginal utility of income is too restrictive leading the log utility model to yield implausible predictions.

The commonly used but ad hoc preference shifter model does a good job of replicating market level average elasticities, marginal costs, pass-through and consumer welfare but is less successful in recovering distributional aspects of demand and welfare effects. If researchers are interested in the distributional consequences of reforms that result in price changes that are small relative to total income, they should consider either flexibly incorporating income effects, or using more flexible models of preference shifting.

In our application the marginal utility of income is clearly non-constant. However, because we consider a small market share good, the change in price induced by the tax is small relative to

$y - p_j$. The policy change itself induces a small income effect. This is important in understanding why the preference shifter model successfully recovers the aggregate consumer welfare change. Similarly, because we find that the curvature of household level demands under the polynomial utility model is similar to that in the more restrictive models in which utility is linear in price, the preference shifter model is able to recover the same pattern of pass-through as the more general model. In applications in which a tax induces a price change that is large relative to $y - p_j$, or in which the curvature of individual demands is less well captured by the log-concave shape of a logit model with utility linear in price, the preference shifter model would do less well at replicating the results of the polynomial utility specification.

In applications to product categories comprising large shares of consumers' budgets, flexibly modeling income effects is likely to be even more important than in our application. In such markets, price changes are more likely to be large enough to induce significant income effects. In applications involving large budget share items (e.g. cars) it has been common to allow for income effects through use of the log utility formulation. Our results suggest this specification may be overly restrictive and insufficiently flexible to capture the true variation in the marginal utility of income and should be tested against more flexible alternative specifications.

Chapter 4

Ownership of intellectual property and corporate taxation

4.1 Introduction

The growing importance of intellectual property as a factor in production,[1] and concern that it is easier for firms to shift income from this source than it is from others, presents challenges for tax design. Firms can and do position their intellectual property with a view to reducing tax liabilities. However, despite these concerns, firms do not by and large locate the legal ownership of intellectual property in the lowest tax countries, and corporate income taxes still raise considerable amounts of revenue in most developed countries. In this paper we address the question of how influential corporate income taxes are in determining where firms choose to legally register ownership of an important form of intangible assets, patents.

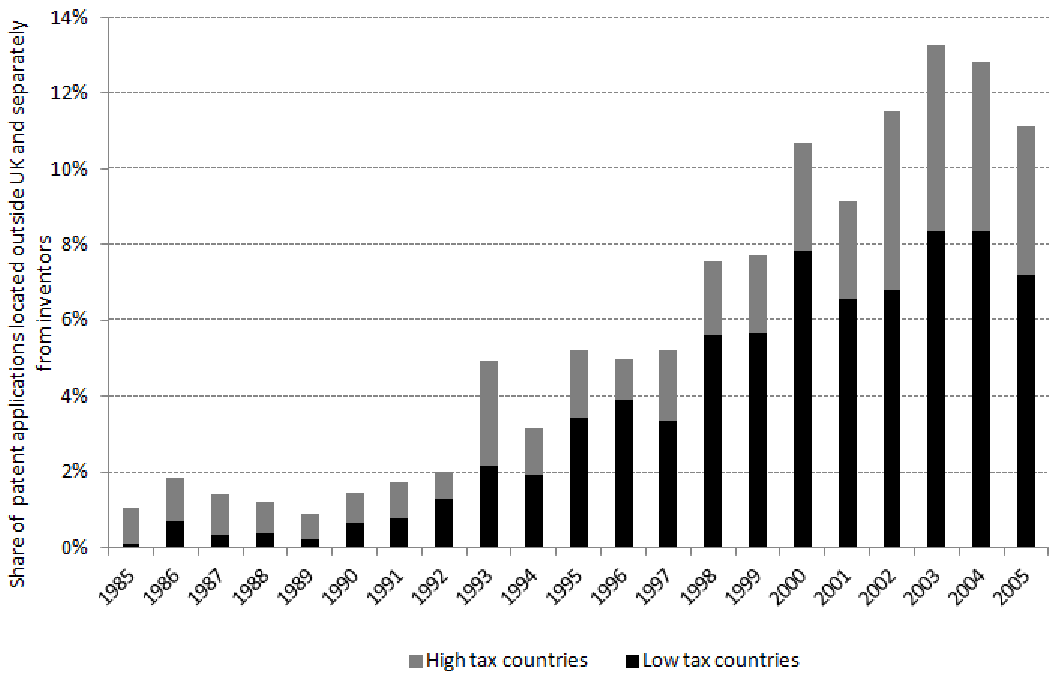
Our contribution is to extend the empirical literature on public policy and firm location choice by introducing new methods to this area of public economics. We estimate a mixed (or random coefficients) logit model that incorporates both observed and unobserved heterogeneity in firms' location choices (see inter alia, Berry et al. (1995, 2004), Nevo (2001) and Train (2003)). A key strength of this approach is that it allows us to compute own and cross tax elasticities across locations that reflect patterns of correlation in observed choices in the data, and therefore to capture more realistic substitution patterns than standard logit models. Our estimates allow us to conduct *ex ante* analysis of how the location of ownership of intellectual property will respond to changes in policy. We use our estimates to simulate responses to recent policy reforms that

[1] OECD (2006) describes the growing significance of intellectual property and its simultaneous use by many different parts of a firm as "one of the most important commercial developments in recent decades". Treasury (2010) estimates that in the UK knowledge investment overtook fixed capital investment in the mid-1990s and is now about 50% higher.

provide preferential tax treatment to income arising from patents. We find that these reforms are likely to have significant effects on the location of ownership of new intellectual property, and could lead to substantial reductions in tax revenue. Our estimates could be used to simulate a wide range of other counterfactual situations.

We use comprehensive panel data on all patent applications made to the European Patent Office (EPO) by a large number of innovative European firms over 1985-2005. A patent is a legal document that grants a firm the exclusive rights to use or licence a novel technology for a specified period of time. A firm can register legal ownership of a patent in a subsidiary that is located in a country different to the firm's headquarters, different to the location where the underlying technology was created and different to the location where the intellectual property will be applied. Lipsey (2010) notes that, in multinational firms, intangible assets *"have no clear geographical location, but only a nominal location determined by the parent company's tax or legal strategies."* For example, Figure 4.1 shows the share of patent applications made by UK parent firms where the legal ownership is registered outside of the UK and in a separate place to where the underlying innovative activity occurred. This share has increased six-fold over the past two decades. The largest proportion has gone to countries that have a lower tax rate than the UK, but the amount going to countries with a higher tax rate has also increased.

Figure 4.1: *Share of patent applications made by subsidiaries of UK parent firms that are located offshore and separately from innovative activity*



Notes: The bars show the share of total patent applications made by subsidiaries of UK parent firms where the subsidiary is located outside of the UK, and is not in a country where associated innovative activity was carried out. Low (high) tax countries are defined as those locations that have a statutory tax rate less (greater) than the UK.

We model the impact of tax on where firms choose to locate the legal ownership of patents. Tax could influence this decision because the legal ownership of the patent will be one of the determinants of where the income derived from the patent is taxed. The profits earned from the exploitation of intellectual property will be the result of a number of activities, including the research and development (R&D) investment undertaken to create the new idea, the financing of this investment and the subsequent commercialisation. When these activities take place in multiple countries, as is often the case for multinational firms, the returns must be allocated to individual jurisdictions for tax purposes.

Firms have an incentive to arrange their activities in such a way that, all else equal, profits accrue in the country in which they would pay the lowest tax. There are a number of strategies that can be used to achieve this. Such strategies commonly require that the income earned from exploiting intellectual property accrues outside of the country in which the underlying R&D took place. One way to achieve this is through contract R&D. For example, a subsidiary in a relatively low tax country may finance (and bear the risk for) R&D activities that are contracted to a related subsidiary in a higher tax country (possibly with the benefit of R&D tax incentives and access to high skills levels). The contract will specify the payment to be made for the R&D activities (commonly equal to the costs incurred plus an arm's length mark-up). Returns above this payment, either from using the technology directly or from licensing it, will accrue to the subsidiary that bore the financial risk. There is a tax advantage to this strategy if the true value of the R&D activities is greater than the price paid for the contract R&D. A similar result may be achieved through the use of a cost sharing agreement that specifies how subsidiaries will share the costs, risks and returns associated with an R&D project. Such agreements may be designed such that the right to exploit and capture the returns from a technology accrues to a subsidiary in a low tax country. The strategies available to a firm depend on how the firm is organized and on the precise tax rules they are subject to (Finnerty and Russo (2007)).

Tax rules limit a firm's ability to manipulate where income arises for tax purposes. Shifting income typically requires that payments made to compensate the company that conducts the R&D, or royalties made for the use of a technology, are at preferential prices. There are transfer pricing rules that aim to enforce the principle that the prices of intra-firm transactions are set as if they had occurred between unrelated parties - this is the arm's length principle. However, these transactions often do not have market counterparts, which means that firms may have opportunities to set the prices of related transactions in such a way as to reduce tax liability.[2] Tax rules, including those that dictate how a firm can allocate the returns to innovative activities,

[2]When determining the correct transfer price there are both conceptual difficulties - it can be hard to separately determine the value that arises from integrated activities that take place across countries - and practical difficulties - firms have more information than tax authorities and an incentive to minimize their tax liability.

differ across European countries and are different to those faced by US multinationals. For example, countries differ on the acceptable methods used to calculate payments for contracted R&D services, and where there are cost sharing agreements, countries differ in the requirements over whether all subsidiaries involved in the agreement need be engaged in R&D (in contrast to the US, not all European countries allow holding companies in low tax locations to be part of cost sharing agreements).

The corporate tax rate is likely to be an important determinant of the location in which a firm chooses to hold legal ownership of intellectual property. However, it is unlikely to be the only factor; we would not expect all intellectual property to be legally registered in the lowest tax countries. Indeed, legal ownership of patents is rarely in the set of small countries that are often considered to be tax havens. The patents that are legally owned in such countries accounted for fewer than 0.5% of all patent applications made to the European Patent Office over the period 2001-2005, and many of those are unrelated to European firms.[3] This could be due, at least in part, to the operation of Controlled Foreign Company (CFC) regimes, which effectively seeks to tax income at the higher home country tax rate if it is deemed to be located in a low tax country for tax purposes. More generally, there may be characteristics of a location over and above its corporate tax rate that firms value. For example, the strength of intellectual property rights protection and market size might play a role, and, all else equal, firms may be more likely to co-locate ownership of intellectual property with associated real innovative activity due to externalities from co-location.

There is likely to be a large degree of heterogeneity in how responsive firms are to tax when deciding where to locate the legal ownership of their intellectual property; a number of papers have emphasized the importance of incorporating heterogeneity in firms' decisions (Melitz (2003), Bernard et al. (2007a, 2007b), Krautheim and Schmidt-Eisenlohr (2011)). This heterogeneity will arise for a number of reasons, some of which relate to observable factors, and others that relate to factors unobserved by the econometrician. For example, firms are likely to be more sensitive to tax when choosing the location in which to legally own patents with a relatively high expected value (Becker and Fuest (2007), Bohm and Riedel (2012)). Firms are also likely to be differentially responsive to tax due to differences in their organisational structures. Their existing network of subsidiaries, the proficiency of their tax department and the tax strategies they are able to employ for managing income from intellectual property will play a role. Firms with headquarters in different countries might respond differently if countries differ in the stringency of their tax rules and in the effectiveness with which they are applied. Firms operating in some markets or using certain technologies might respond differently, because, for example, transfer

[3]Figure based on patent applications made by applicants located in Bahamas, Barbados, Bermuda, Cayman Islands, Gibraltar, Hong Kong, Liechtenstein, Malta, Monaco, Netherlands Antilles, Panama or Singapore.

pricing rules may be easier to circumvent for firms operating in markets where a high share of transactions are intra-firm meaning it is difficult for tax authorities to accurately assess what is a fair market price. Both firm size and industry have been highlighted as important in the context of firm decision making over how to organize offshore activities (Graham and Tucker (2006) and Desai et al. (2006)). Indeed, the value of a patent, the relative attractiveness of a location and a firm's strategies and organisational structures are likely to vary across industries and, within industries, across firms.

Our work relates to several papers in the literature. Most closely related, Dischinger and Riedel (2011) and Karkinsky and Riedel (2012) estimate the relationship between corporate tax and, respectively, the quantity of intangible assets and the number of patent applications made by subsidiaries located in each of a number of European countries. Also related is Ernst and Spengel (2011) who estimate the impact of R&D tax incentives and corporate tax on patenting. In common with these papers, we are interested in the relationship between corporate tax and where firms choose to locate intellectual property. We extend this literature by estimating a choice model that allows us to compute the full set of own tax and cross tax elasticities and which allows us to carry out *ex ante* analysis of how location decisions will respond to potential policy changes. Our work is also related to Cohen (2012), which uses a discrete choice framework to study how the design of US state tax rules influence US firms' decisions over in which state to incorporate.

There is a considerable literature in the Hall and Jorgenson (1967) tradition that considers the impact of taxes on production activity and on the location of R&D. Hines (1996, 1999) and Devereux and Maffini (2006) provide surveys of the empirical literature. This literature finds that, despite the many factors that will influence a firm's location decision, tax exerts a significant effect on location choices. Hines and Jaffe (2000) show that tax affects the location of firms' innovative activities within US multinational groups. Most relevant for our analysis, previous work has highlighted the role that intangible assets play in allowing firms to organize their activities with a view to reducing their tax burden (Grubert and Altshuler (2008)). Empirical studies provide indirect evidence of tax avoidance by, for example showing that firms have relatively high profitability in low tax countries (Grubert and Mutti (1991), Hines and Rice (1994)) and that the share of royalty payments associated with low tax countries is higher than expected (Mutti and Grubert (2009)). Grubert (2003) formalizes how intangible assets can be used to shift income and finds that about half of the income shifted from high-tax to low-tax countries by US manufacturing firms can be accounted for by income from R&D linked intangibles.

The structure of the paper is as follows. In Section 4.2 we outline a model of a firm's decision over where to locate the legal ownership of a patent. In Section 4.3 we describe the data we use to estimate the model. Section 4.4 presents the estimated coefficients and the tax elasticities between locations. An example of how the model can be used to conduct policy simulations is given in Section 4.5, where we consider the impact of recent reforms that reduce the tax rate for income derived from patents. A final section summarizes and concludes.

4.2 Firm behavior

When a firm generates a new idea, it expects to earn a stream of income on the application of that idea in the future. Ideas will vary both in their expected values and in the number of patents they give rise to (some will lead to one patent and some will lead to many). A firm faces the decision over where to initially locate the legal ownership of each patent. It will make this decision based, in part, on the rate of tax that it expects to face on income generated by the use of the patent in the future. Unobserved attributes of ideas are likely to be crucial, and potentially could generate correlations in patent location decisions. The firm will also take account of other characteristics of locations that it may value, for example, whether the real innovative activity associated with that intellectual property is also located there, the potential size of the market (if it also expects to commercialize the idea in that location), intellectual property rights' protection, technological condition, and many other location specific factors, at least some of which are likely to be unobserved by the econometrician. The importance of these location characteristic are likely to vary across ideas. For example, high value ideas may be more tax sensitive and the importance of intellectual property protection may differ across industries.

We develop a tractable empirical model that captures these determinants of location choice.

4.2.1 Firm payoffs

We specify a model in which a parent firm decides where to locate the legal ownership of each of its patents. Firms, indexed $f = 1, ..., F$, realize ideas, indexed $i = 1, ..., I$. Ideas are assumed to arise exogenously over time, indexed t. Each idea can yield a single patent or a group of related patents; patents are indexed $p = 1, ..., P$. We model the country, indexed $j = 1, ..., J$, in which the parent firm decides to locate the legal ownership of each patent, allowing for correlation in decisions between related patents (those that are part of the same idea). We consider all patents taken out by a parent firm that are technologically related in a quarter as part of the same idea; the precise definition of an idea is given in Section 4.3.1.

For each patent, the parent firm chooses the location that yields the highest payoff. The payoff the parent firm gets from choosing a location depends on the tax rate it expects to face, τ_{fjt}, the quality of the idea, q_i, whether any research activity that gave rise to the idea is located there, a_{ijt}, the strength of the country's intellectual property rights protection, the size of the local market (measured as GDP), and the level of technological innovativeness (measured as total annual business R&D expenditure as a share of GDP), captured in the vector x_{jt}. Crucially, we also allow location choice to depend on unobserved characteristics of both the idea and the location. We allow the impact of all observed and unobserved factors to vary across medium and large firms and for technologies in different industries; the subscript $r = 1, ..., R$ denotes the industry-firm size category an idea belongs to.

We assume the payoff that firm f obtains from placing legal ownership of patent p (belonging to idea i) in location j takes the form,

$$\pi_{pjt} = \alpha_i \tau_{fjt} + \beta_i a_{ijt} + \gamma_r x_{jt} + \xi_{rj} + \epsilon_{pjt}. \tag{4.1}$$

The parameters α_i and β_i vary across ideas and are functions of both observable and unobservable idea characteristics,

$$\alpha_i = \bar{\alpha}_r + \alpha_r q_i + \sigma_r^\tau \eta_i;\ \eta_i \sim \mathcal{N}(0, 1), \tag{4.2}$$

$$\beta_i = \bar{\beta}_r + \sigma_r^a \nu_i;\ \nu_i \sim \mathcal{N}(0, 1). \tag{4.3}$$

We assume that η_i and ν_i are uncorrelated with each other and with the other covariates, and that the additive shock ϵ_{pjt} is distributed iid type I extreme value. ξ_{rj} is a location-industry-firm size fixed effect. The firm chooses option j^* if,

$$\pi_{pj^*t} > \pi_{pjt}\ \forall\ j \neq j^*. \tag{4.4}$$

The resulting choice model is a mixed logit with unknown parameters $(\bar{\alpha}_r, \alpha_r, \sigma_r^\tau, \bar{\beta}_r, \sigma_r^a, \gamma_r, \xi_{rj})$.

The tax rate τ_{fjt} varies across firms, because the tax system in a firm's residence jurisdiction may interact with the rules of the countries in which it is considering locating ownership of a patent through the operation of Controlled Foreign Company (CFC) rules. We use the tax rate dated t, making the assumption that when a firm chooses the location of a patent it expects that the current tax regime will apply in the future.

The tax parameter, α_i, is at the idea level. We allow it to vary with an observed measure of idea quality, q_i. Patents that are part of a high quality idea are likely to have a higher expected value, and thus their location may be more sensitive to tax. We also allow the idea level tax parameter to include a random term η_i. This captures all components of ideas that determine the responsiveness of location choice to tax and are observed by the parent firm but not by the econometrician. For instance, the quality variable is likely to be an imperfect measure of the expected value of the idea. There could be other factors that are correlated with the idea's expected values, unobserved by us, but available to firms, which will be captured by η_i.

The parameter $\bar{\alpha}_r$ captures the mean marginal effect of tax on the payoff, α_r tells us how this varies with the observed quality of the idea, and σ_r^τ tells us the standard deviation in the effect of tax on the payoff. These parameters all have an r subscript, indicating that we allow both the mean impact of tax on the payoff, and how this varies with both observed quality and unobservables, to vary across industries and across firm size.

Similarly, we model the parameter on real innovative activity, β_i, as an idea level random coefficient; the impact of real innovative activity on the payoff function varies across patents with the random term ν_i. Firms may value locating the legal ownership of intellectual property in the same country as it was created, and the strength of this motive is likely to vary across ideas.

A central assumption of the standard multinomial logit model is that the stochastic error term associated with the payoff from a particular option (in our case the decision to locate ownership of a patent in a particular location) has an iid type I extreme value distribution. This rules out correlation in latent payoffs. This assumption leads to a closed form solution for the location choice probabilities, which is empirically convenient. However, it is restrictive, leading the multinomial logit model to imply restrictive substitution patterns. In particular, the lack of correlation in payoffs endows the model with the independence of irrelevant alternatives (IIA) property.

Random coefficients allow us to relax the assumption of zero correlation in payoffs inherent in the standard multinomial logit model. In particular, we group patents into ideas and allow some of the parameters in the payoff function (those on tax and on real innovative activity) to be idea specific. We model these idea specific preference parameters as random coefficients - which leads to the mixed logit model. An implication of including random coefficients at the level of an idea is that it allows for correlation in the payoffs both across location options for a given patent, and across patent locating decisions for all patents in a given idea (Train (2003)); the covariance between the payoff of locating patent p in location j and locating another patent p' associated with the same idea in location k is given by: $\text{Cov}(\pi_{pjt}, \pi_{p'kt}) = \sigma_r^\tau \tau_{fjt}\tau_{fkt} + \sigma_r^a a_{ijt} a_{ikt}$. Therefore, although the model still includes an iid type I extreme value term, it also contains other unobservable components that allow for rich correlations, the nature of which is in large part determined by the data (along with the distributional assumptions we make). An important consequence of allowing for such correlations is that the IIA property is not present in the mixed logit model, allowing the model to capture much more flexible substitution patterns (see Berry et al. (1995, 2004), Nevo (2001) and Train (2003)).

We include in x_{jt} a number of time varying location characteristics that firms are likely to value when choosing where to locate the legal ownership of intellectual property. However, there are likely to be other location characteristics that firms value that we do not observe. To capture these we include location-industry-firm size fixed effects, ξ_{rj}. These will control for location specific costs, such as the legal costs associated with setting up a subsidiary, or location specific benefits, such as government provided public goods, the relevance of which might differ for firms of different size or industry.

Note that there may be many other patent, idea or firm specific factors that do not vary across location but that do influence the costs or benefits of a location. However, because these do not vary across location they will not enter the location choice decision, and therefore are not explicitly entered into equation (4.1); including them would lead to an observationally equivalent choice model, because they would drop out when payoff comparisons are made across locations.

4.2.2 Choice probabilities

The choice model described above implies that the probability that legal ownership of patent p is located in location j, conditional on realisations of the idea specific variables η_i and ν_i, takes the form,

$$\rho_{pjt}(\eta_i, \nu_i) = \frac{e^{(\bar{\alpha}_r+\alpha_r q_i+\sigma_r^\tau \eta_i)\tau_{fjt}+(\bar{\beta}_r+\sigma_r^a \nu_i)a_{ijt}+\gamma_r x_{jt}+\xi_{rj}}}{\sum_k e^{(\bar{\alpha}_r+\alpha_r q_i+\sigma_r^\tau \eta_i)\tau_{fkt}+(\bar{\beta}_r+\sigma_r^a \nu_i)a_{ikt}+\gamma_r x_{kt}+\xi_{rk}}}. \tag{4.5}$$

The unconditional probability is obtained by integrating out the unobservable idea specific random terms,

$$P_{pjt} = \int \rho_{pjt}(\eta_i, \nu_i) dF(\eta_i, \nu_i). \tag{4.6}$$

Equation (4.6) can be used to compute the impact of marginal changes in tax on location choice probabilities. For instance, the elasticity of the probability that legal ownership of patent p is in location j with respect to a marginal change in the tax rate in location k is given by,

$$e_{pjkt} = \frac{\partial P_{pjt}}{\partial \tau_{ikt}} \frac{\tau_{ikt}}{P_{pjt}} = \int \frac{\partial \rho_{pjt}(\eta_i, \nu_i)}{\partial \tau_{ikt}} dF(\eta_i, \nu_i) \frac{\tau_{ikt}}{\int \rho_{pjt}(\eta_i, \nu_i) dF(\eta_i, \nu_i)}. \tag{4.7}$$

We compute the elasticity of the share of patents with legal ownership in location j with respect to the statutory tax rate in location k in year t (τ_{kt}), by aggregating across equation (4.7) for all the patents that arose in year t (denote this set of patents Υ_t). We explicitly account for the operation of CFC regimes (see Section 4.3.2 for a description of how CFC regimes work).

For a patent owned by a firm resident in a country that has a CFC regime, and that deems country k as a low tax jurisdiction, changes in the statutory tax rate of location k should have no bearing on their location probabilities. Define an indicator variable D_{pjt}, where $D_{pjt} = 0$ if patent p at time t is subject to CFC rules which bind in location j, and $D_{pjt} = 1$ otherwise. We can then write the elasticity of the share of patents located in country j with respect to tax in location k in year t as,

$$E_{jkt} = \int \sum_{p \in \Upsilon_t} D_{pkt} \frac{\partial \rho_{pjt}(\eta_i, \nu_i)}{\partial \tau_{ikt}} dF(\eta_i, \nu_i) \frac{\tau_{kt}}{\int \sum_{p \in \Upsilon_t} \rho_{pjt}(\eta_i, \nu_i) dF(\eta_i, \nu_i)}. \tag{4.8}$$

4.2.3 Identification

Our primary interest lies in pinning down the *ceteris paribus* impact of a change in the corporate tax rate set by any one country on the shares of patent applications made by subsidiaries in both that and in alternative locations. To do this we must consistently estimate the parameters of the payoff function outlined in equations (4.1)-(4.4), and in particular the parameters governing the marginal impact of tax on the payoff associated with selecting a location, which are modelled as random coefficients. Train (2003) shows that the random coefficient model has a duel interpretation as an error component model. Under this representation, the mean of the random coefficient can be interpreted as a fixed coefficient - it is pinned down by variation in location choices in response to variation in taxes faced by firms, conditional on other observables included in the model. The standard deviation of the random coefficient is interpreted as a component in the error term - it is pinned down by correlation in payoffs across locations (both in a given choice set and across location decisions within a given idea).

Berry and Haile (2010) have established that in random utility multinomial logit models the distribution of unobserved preference parameters is non-parametrically identified given sufficiently rich micro data. However, non-parametric estimation of this model is computationally burdensome, and we therefore follow most papers in the literature by assuming payoffs are linear with independent additive shocks and that the distribution of unobserved parameters is normally distributed. These assumptions give us a convenient approximation.

The standard identification concerns still apply here; to consistently estimate the parameters we require that the additive shock (ϵ_{pjt}) and the idea specific random terms (η_i and ν_i) are independent from each other and from the other explanatory variables. Specifically, if there are factors which influence location choice, and that are not captured by the observed and unobserved controls that we include, this would lead to inconsistent estimates. To mitigate this concern, we include a number of controls in the model. We include location-industry-firm size fixed effects; these control for all country characteristics that affect a firm's payoff and that do not vary through time, but that potentially do vary across firms in different industries and different parent firm sizes.

We include time varying (non-tax) location characteristics. These include a measure of the presence of real innovative activity associated with the intellectual property. This controls for the fact that some firms, for reasons other than the tax rate they will face in a particular location, may wish to co-locate legal ownership of intellectual property with real innovative activity. If decisions over the location of real innovative activity are influenced by corporate tax rates, then failure to control for this would result in an inconsistent estimate of the impact of tax on patent location choice. We also control for the strength of intellectual property rights protection in the location, market size and technological innovativeness - all factors that vary over time and location, and may be expected to impact on intellectual property location choices.

Identification of the tax coefficients also relies on the presence of informative variation in taxes in the data; specifically we need to observe variation in the set of taxes in potential locations across patent choice situations, conditional on all other factors that influence location choice. Crucially, it is necessary that there is variation in *differences* in taxes between locations across choice sets. So, for instance, if the only source of tax variation was that all tax rates changed simultaneously by the same amount, the marginal effect of tax on the payoffs would not be identified.

Such variation arises in our framework for two reasons. First, there is variation over time in statutory tax rates; as outlined below in Section 4.3.2, over the period for which we have data (1985-2005) there has been a general pattern of declining statutory tax rates. The size of this decline has varied across countries, and the changes have occurred at different times, meaning that tax reforms have given rise to variation in the set of tax rates across locations. Second, in addition to this time series variation, CFC regimes lead to another source of variation; two parent firms taking a decision at the same point in time, but resident in different countries, can face a different set of tax rates due to cross country differences in CFC rules. The variation in location choices, conditional on other factors, in response to this variation in tax rates pins down the impact of tax on location choice.

4.3 Data

To estimate the model we need information on where firms have chosen to locate the legal ownership of their patents, the corporate tax regime and other conditioning variables.

4.3.1 Patents data

We use data on patent applications filed at the European Patent Office (EPO) by the European and US subsidiaries of parent firms located in fourteen European countries. We exclude from our analysis firms that patent infrequently. The number of patent applications by location of the subsidiary that filed the patent application is shown in Table 3.1. Our data include 1083 parent firms that collectively have 4,823 patenting subsidiaries, which file 379,849 patent applications over the period 1985-2005. These account for a 70% of all corporate applications filed at the EPO by firms parented in these fourteen European countries during this period.

Each patent application lists the firm that files the application (the applicant), this is the legal owner of the patent.[4] We identify the parent firm using information from company accounts (from Amadeus), company websites, business directories and other sources (see Abramovsky et al. (2008) for details). We use ownership information at a fixed point in time (2004), and we do not observe changes in ownership after an application has been filed.

For each patent application this gives us a mapping between the location of the parent firm and the location of the subsidiary that legally owns the patent. We also observe the location of the inventors (individuals) that created the technology underlying the patent application.[5] There are often inventors located in multiple countries, and often in different countries to that of the applicant.[6] The location of both the applicant and the inventors are distinct from the patent office to which the firm is applying for protection. For patent applications filed at the EPO, each application also designates the individual countries in which final patent protection will be sought; it is these individual countries, not the EPO, that grants patent protection.

We use Thomson's Derwent database to classify patent applications based on the technology embodied in the patent and the markets in which the technology is used. We use three broad industry groups - Chemical, Electrical and Engineering. A patent application can be relevant for more than one industry group if it has applicability in more than one of these industries. Where this is the case we include the patent application in each of the industry sub-samples, and when we calculate the market level elasticities we weight each patent application so that the sum of the weights equals one (so if a patent application is in two industries it will get a weight of 0.5 in each). Table 4.1 columns (2)-(4) shows the industry split of patent applications, 32.3% are in Chemical, 36.8% in Electrical and 30.8% in Engineering, but this varies across countries. We restrict our analysis to firms that are above the 20th percentile in terms of the number of

[4]In most EPO applications there is one applicant. For the handful of patent applications that are filed by multiple applicants in multiple locations we randomly select one of them.

[5]A small number of patent applications (0.07% of those filed by all applicants in the 15 countries we consider) have missing inventor data. We exclude these applications.

[6]This is different to data on patents filed at the US Patent and Trademark Office for which the inventors are the patent applicants.

patent applications per firm in their industry. We distinguish large firms as those with a level of patenting above the 80th percentile in their industry. 78.8% of patent applications in our sample are held by large firms.

Firms often take out a number of related patent applications at the same time; we allow for correlation in these decisions. We group together related patent applications that can be considered to be part of the same idea. We identify patent applications as part of the same idea if they are made by the same parent firm, are filed in the same quarter (i.e. three month period), are classified in the same industry and share a network of common inventors. The number of patent applications in an idea vary: on average an idea contains one patent application; ideas containing more than one application account for 26% of all patent applications.

The importance of ideas in our empirical strategy is that we allow correlation across patent applications at the level of the idea; the decisions over where to locate ownership of these related patent applications is unlikely to be independent, and the inclusion of random coefficients at this level allows for them to be correlated.

Patent quality

There is a large literature that highlights the skewness of patent value and quality (Pakes (1985, 1986), Blundell et al. (1999), Lanjouw and Schankerman (2004) and Hall et al. (2007)). Firms file patent applications for a variety of reasons; some are filed to protect valuable new ideas, others are filed strategically to provide option values or to block competitors. Patent value also varies because some ideas are commercially more valuable than others. We identify high quality patent applications as those that are part of a triadic patent family, i.e. a related patent application has been filed at each of the EPO, the US Patent and Trademark office and the Japan Patent Office. The OECD uses triadic patent families *"... to improve the international comparability and quality of patent-based indicators ... patents included in the triadic family are typically of higher economic value: patentees only take on the additional costs and delays of extending the protection of their invention to other countries if they deem it worthwhile."* (OECD (2012))

Table 4.1: *Patent applications and tax rates by applicant country*

	Number	% patent applications:									Number
	of patent	by industry			by firm size		high	Statutory tax rate			of rate
Applicant country	applications	Chemical	Electrical	Engineering	Large	Medium	quality	min	max	mean	changes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Belgium	6935	47.8	22.9	29.3	74.2	25.8	29.0	34.0	45.0	40.4	5
Denmark	4975	56.9	15.2	27.9	65.7	34.3	31.2	28.0	50.0	36.0	7
Finland	10151	13.9	72.6	13.5	83.1	16.9	33.8	25.0	60.2	36.8	7
France	50017	26.6	45.2	28.2	87.5	12.5	37.1	33.3	50.0	38.6	13
Germany	166606	29.1	34.3	36.6	78.5	21.5	35.8	38.3	61.7	52.5	15
Ireland	387	70.6	7.7	21.7	52.5	47.5	26.0	10.0	12.5	11.3	1
Italy	11155	34.1	26.3	39.5	47.9	52.1	29.9	37.3	53.2	44.6	7
Luxembourg	792	52.9	25.6	21.4	32.8	67.2	41.4	30.4	39.4	37.5	6
Netherlands	34918	29.3	51.4	19.3	90.2	9.8	36.1	31.5	43.0	37.2	4
Norway	1083	35.3	29.4	35.3	47.7	52.3	36.8	28.0	50.8	39.4	1
Spain	994	44.7	19.8	35.5	23.9	76.1	33.5	35.0	35.0	35.0	1
Sweden	17300	21.1	49.0	29.9	82.7	17.3	43.3	28.0	52.0	36.7	2
Switzerland	30221	45.6	23.7	30.7	73.8	26.2	39.6	21.3	28.5	25.1	6
United Kingdom	27955	45.0	30.0	25.0	67.3	32.7	38.2	30.0	40.0	33.8	5
United States	16360	48.7	29.9	21.4	88.6	11.4	37.9	34.0	49.5	40.1	5
Total	379849	32.3	36.8	30.8	78.8	21.2	36.5	10.0	61.7	43.2	6

Notes: Applicant country refers to the location of the subsidiary that is the applicant on a patent application. The final row gives means across all patent applications in our data. The mean tax rate in the final row is averaged across all choice options, so differs from Table 3.3.

We expect triadic patent applications to be of a higher value since there is a cost to filing patent applications at each of these patent offices, and the main incentive to do this is if firms expect the technology to have a wide application. Each idea (group of patent applications) is classified as high quality if over half of the associated patent applications are triadic. As seen in Table 4.1 (column (7)) on average 36.5% of patent applications are classified as being part of a high quality idea.

Patent ownership and income from patents

We model the impact of tax on where firms choose to locate the legal ownership of patents. In the introduction we discuss the reasons that we might expect tax to affect a firm's decision of where to hold legal ownership of intellectual property.

The extent to which firms have arranged their activities in such a way that income can reasonably be deemed to be attributable to the subsidiary that legally owns the intellectual property will differ; firms will differ in how aggressively they seek to manage their tax liabilities. For some firms, the choice of where to earn income may be a choice between those countries in which real innovative activities already takes place; others may employ strategies that allow them to earn income in a separate country.

There are many factors that effect the costs and benefits of choosing a particular location. For tax havens these costs may be particularly high: CFC rules are more likely to bind; the transfer of profits to locations where there is little real activity will be more difficult; tax havens are likely to be less attractive locations along non-tax dimensions such as intellectual property rights protection. We would not expect legal ownership of all patents to be located in such countries. However, it is possible that some firms are particularly aggressive in their tax planning and organize their activities in such a way that income is earned in a location that is not where legal ownership is located. We do not observe income flows, so we do not explicitly model this behavior; to the extent that it makes the decision over the location of legal ownership less related to tax we would be less likely to find an impact of tax. In our model we aim to capture this variation in behavior across firms, and within firms across ideas, by the inclusion of observed and unobserved heterogeneity.

An additional complicating factor is that a firm might file a patent application from one subsidiary, but later transfer ownership of that patent to another related firm. However, firms have an incentive to consider tax when making the initial location decision, because in many situations there are tax costs to transferring the ownership of intangible assets. For there to be a tax benefit to the sale or transfer of an asset it must be the case that this can happen at a value below the true market value. The transfer of intellectual property will be subject to transfer pricing rules, which will act to limit how much value can be shifted to a low tax country. In addition, many European countries operate exit taxes that attempt to levy tax on the net present value of the expected revenue stream on an intangible asset when it is moved out of the country. Such tax provisions reduce (if not remove) any tax advantages to re-locating to a lower tax jurisdiction. If firms do intend, with some positive probability, to re-locate the ownership of a patent in the future, and if transfer pricing rules and exit taxes do not act to perfectly off-set any tax advantages of doing so, this would reduce the importance of corporate tax in the initial location decision. This is an additional reason that it is important that we allow heterogeneity in the importance of corporate tax across intellectual property.

The place where we need to make more restrictive assumptions about the relationship between legal ownership and income from intellectual property is when we carry out the *ex ante* analysis of the Patent Box tax reforms and calculate the revenue implications of these reforms. In order to do this we need to assume that the relationship between legal ownership and income is not changed by the policy reform.

4.3.2 Taxes

We measure the impact of tax on payoffs using the statutory tax rate. We assume that returns from intellectual property are expected to be sufficiently high that deductions such as capital allowances are relatively unimportant, so that the effective tax rate faced by the firm is approximately the statutory tax rate (see Devereux and Maffini (2006), where Figure 1 shows that the marginal effective tax rate asymptotes to the statutory tax rate as profitability increases).

Our identification strategy relies on variation over time and across countries in the tax rate. Table 4.1 (columns (8)-(11)) summarise the variation in corporate tax rates. In general, main statutory tax rates fell in the two decades up to 2005, but with the timings of changes differing across countries. The Scandinavian countries - Denmark, Finland, Norway and Sweden - reduced tax rates significantly around 1990. Italy enacted a reduction of over 10 percentage points in 1998, as did Germany in 2001. France and the UK have enacted a series of gradual reductions.

Table 4.2: *Number of firms and CFC regimes by parent firm country*

Patent firm country	Number of parent firms	CFC regime introduced	Applicant countries for which CFC ever binds (no. of years*)
Belgium	28	-	-
Denmark	27	1995	FI (1 year), IE (all years)
Finland	25	1995	IE (11 years)
France	108	1980	CH (6 years), IE (all years)
Germany	446	1972	CH (all years), FI (8 years), GB (1 year), IE (all years), NO (9), SE (10)
Ireland	4	-	-
Italy	75	2002	-
Luxembourg	8	-	-
Netherlands	58	-	-
Norway	7	1992	IE (14 years)
Spain	8	1996	CH (7 years), FI (2 years), IE (all years)
Sweden	48	1990	CH (1 year). IE (11 years)
Switzerland	106	-	-
United Kingdom	135	1984	CH (2 years), IE (all years)

Notes: Country codes: Belgium (BE); Switzerland (CH); Denmark (DK); Finland (FI); France (FR); Germany (DE); Ireland (IE); Italy (IT); Luxembourg (LU); Netherlands (NL); Norway (NO); Spain (ES); Sweden (SE); United Kingdom (GB)

There can be additional tax levied in the parent firm's home country as a result of Controlled Foreign Companies (CFC) rules, which aim to prevent firms locating income in lower tax countries in order to avoid taxation in their home country. CFC rules set out criteria for identifying subsidiaries that are located in a country deemed to be 'low-tax' and earning a significant amount of 'passive income' (income that is not associated with real activity). When a CFC regime is in place in a parent firm's country of residence, and a subsidiary is located in a country that is deemed a 'low tax' location (as judged against parent firm country specific thresholds), then we set the tax variable, τ_{fjt}, equal to the parent firm country's statutory rate. A description of the country pairs for which this is the case is given in Table 4.2. There is variation in whether a parent firm country operates a CFC regime (some regimes are introduced during the period for which we have data) and in the applicant countries that are deemed low tax (which differ over time when statutory rates change). This definition of whether CFC rules bind effectively assumes that the income received from a patent is deemed to be passive income, and that the share of passive income is sufficient to trigger the CFC rules. This is clearly an approximation. However, if we look across all location options that firms in our data face and that are deemed low tax by CFC regimes, then it is rarely the case that the parent firm has both inventors and holds legal ownership of a patent application in the same location. The results we present below are robust to the alternative assumption that patent applications with ownership located in countries where associated real innovative activity is also located would be treated as active income, so that CFC rule do not bind.

4.3.3 Descriptive statistics

The variables included in the model are defined and summarised in Table 4.3. The top panel contains the observed location attributes we include. These comprise the tax rate that the parent firm would face if it earned income from the application of intellectual property in the location, a measure of the presence of real innovative activity in a location defined as an indicator of whether at least one of the inventors associated with the patent applications that form the idea are located in that country and country-time varying observable characteristics. The latter includes a measure of intellectual property rights protection. This is based on a measure developed by Ginarte and Park (1997) and Park (2008). The countries we consider all have advanced systems of property rights and therefore rank relatively highly on the protection of intellectual property. We define a country as having a strong intellectual property regime if it scores above the median of countries in our sample. Other country-time varying variables include market size, as measured by Gross Domestic Product (GDP) and the technological innovativeness of a country, proxied by business R&D investment in the country as a share of GDP.

We allow the valuations firms place on location characteristics to vary across patent applications. A summary of observable patent (or idea) characteristics is given in the bottom panel of Table 4.3. In estimation we allow all coefficients to vary with the industry the patent application belongs to and the size of the associated parent firm. This allows the model to capture, for example, that large firms are more likely to have organisational structures that assist the location of intellectual property for tax purposes. The tax rate is interacted with a measure of the idea quality, reflecting the possibility that firms' location choices may be more responsive to tax when they expect intellectual property to earn higher returns.

Table 4.3: *Variable definitions and summary statistics*

Variable	Definition	Min	Max	Mean	Stand. dev.
Location characteristics					
Tax rate (τ_{fjt})	Statutory tax rate in applicant country; or statutory tax rate in parent firm county when binding CFC regime	10.00	61.70	41.79	11.34
Real activity	Dummy equal to one when any of the inventors associated with the patent applications that form an idea are located in that country	0.00	1.00	0.86	0.35
Strong intellectual property protection	Measure of applicant countries' relative degree of intellectual property rights protection	0.00	1.00	0.76	0.43
Market size	GDP measured in millions of constant PPP US dollars	0.01	12.56	1.90	1.97
Technological innovativeness	Business investment in R&D as a share of GDP	0.29	3.20	1.53	0.40
Patent characteristics					
Large firm	Large parent firms are those for which the total number of patent applications is above the 80th percentile	0.00	1.00	0.79	0.41
High quality	Ideas are classified as high quality if over half of the associated patent applications were filed at each of the EPO, USPTO and JPO	0.00	1.00	0.36	0.48
Electrical	Instrumentation, computer, electronics, communications, electrical	0.00	1.00	0.37	0.42
Chemical	Chemicals, pharmaceuticals, printing, petroleum	0.00	1.00	0.32	0.42
Engineering	General and mechanical engineering	0.00	1.00	0.31	0.38

Notes: Statistics are based on all patent applications in our data. GDP is measured in constant PPP US dollars (expenditure measure) using a 2005 base year. Business Investment in R&D as a share of GDP (BERD) is from OECD Main Science and Technology Indicators. GDP and BERD available at http://stats.oecd.org.

4.4 Results

Table 4.4 shows the estimated coefficients of the choice model outlined in Section 4.2. The model is estimated using simulated maximum likelihood (see Train (2003)). We allow all coefficients to vary across industry and firms size, indicated by the different columns. We include a full set of location-industry-firm size fixed effects (not reported in Table 4.4, but available upon request).

The top row of Table 4.4 shows that the mean marginal impact of tax on the payoff from placing legal ownership of a patent in a location is negative and statistically significant across all industries and parent firm size groups. The second row shows that in both the electrical and engineering industries the payoff for high quality patents is more sensitive to taxes. This is true both for large and medium firms. In the chemical industry the payoff for a high quality patent is estimated to be marginally less responsive to tax than for lower quality patents for large firms, with there being no statistically significant difference between the high and low quality patents for medium firms. Row three shows that there is a substantial degree of unobserved heterogeneity in the importance of tax on location choice across ideas, the standard deviations of the random coefficients on tax are both large and statistically significant across all industries and size categories.

The fourth row shows that, *ceteris paribus*, having real innovative activity in a location is associated with a higher payoff from placing legal ownership of a patent in that location across all industries and size categories; the fifth row shows that there is a significant amount of variation in the importance of this characteristic across ideas.

Together the large and statistically significant standard deviations on the random coefficients on tax and real innovative activity (in all industry-firm size groups) indicates the presence of important correlations in payoffs, both across locations for a given patent, and across patents in a given idea. These correlations will generate patterns of substitution that will depart from the more restrictive patterns implied by a standard multinomial logit model.

The remaining three rows of Table 4.4 describe the impact of having strong intellectual property protection, and the marginal impacts of market size and technological innovativeness, on the payoff function. For five of the six industry-firm size groups, a location having strong intellectual property protection is, all else equal, associated with firms obtaining higher payoffs from locating legal ownership of their patents there (the exception is medium electric firms, for which the strong intellectual property rights dummy is negative). Larger market size is associated with statistically significantly larger payoffs for five of the size industry-firm size groups, and a higher degree of technological innovativeness is associated with statistically significantly larger payoffs for four of the size industry-firm size groups.

Table 4.4: *Estimated parameters*

Industry	Electrical		Engineering		Chemical	
Size	Large	Medium	Large	Medium	Large	Medium
Tax rate (mean)	-3.48	-4.93	-4.91	-4.88	-6.54	-4.03
	(0.13)	(0.32)	(0.14)	(0.24)	(0.13)	(0.28)
Tax rate x quality	-0.67	-1.98	-0.69	-0.66	0.34	0.08
	(0.14)	(0.37)	(0.15)	(0.28)	(0.14)	(0.31)
Tax rate (std. dev.)	3.59	3.67	2.35	3.17	3.36	4.24
	(0.16)	(0.40)	(0.23)	(0.27)	(0.15)	(0.27)
Real activity (mean)	5.60	7.27	6.20	7.03	6.11	6.55
	(0.03)	(0.12)	(0.05)	(0.09)	(0.04)	(0.09)
Real activity (stand. dev)	2.50	2.76	2.80	2.96	2.95	2.59
	(0.03)	(0.10)	(0.04)	(0.08)	(0.04)	(0.08)
High IP property protection	0.63	-0.22	0.28	0.19	0.68	0.60
	(0.04)	(0.11)	(0.06)	(0.10)	(0.06)	(0.10)
GDP	-0.20	0.66	0.53	0.43	0.39	0.20
	(0.02)	(0.07)	(0.03)	(0.05)	(0.02)	(0.05)
R&D expenditure/GDP	0.26	0.03	0.23	0.09	0.18	0.00
	(0.02)	(0.05)	(0.03)	(0.04)	(0.02)	(0.05)

Notes: Dependent variable is location choice in one of the locations shown in Table 3.1. Estimation is based on 379,849 patent applications. Industry-location-firm size fixed effects included. Large firms are those associated with a total number of patent applications above the 80th percentile in each industry.

Table 4.5 shows the matrix of own and cross tax elasticities implied by the choice model. It contains the elasticities of the share of patents located in each of 14 European countries with respect to the rate of corporate tax set in each of these countries and in the US. These are calculated as described in Section 4.2.2. We report the matrix of elasticities using tax rates and the distribution of patent applications for the most recent year in our data, 2005. Each cell shows the elasticity of the share of patents located in the country indicated in column 1 with respect to the tax rate set by the country in row 1. The emboldened diagonal shows the own tax elasticities. For all locations, except Luxembourg, the own tax elasticities are less than one in magnitude.

There is a limited literature on the elasticity of the location of corporate income with respect to tax. De Mooij and Ederveen (2008) report that empirical studies considering the effect of differences in statutory tax rates on various measures of profitability (with a view to indirectly capturing the effects on profit shifting) tend to find a semi-elasticity of around -1.2. As in this paper, Karkinsky and Riedel (2012) consider the link between corporate tax rates and patent applications. They estimate a semi-elasticity that, depending on the functional form of their model, implies that a 1 percentage point increase in the rate of corporate tax translates into a 3.5%-3.8% fall in patent applications from that location. Direct comparison with our results is made difficult by the fact that our model allows tax effects to vary across all locations. We find that the share of patents held in Luxembourg is most sensitive to tax (the Luxembourg semi-elasticity is 3.9%) and least sensitive for Germany (the German semi-elasticity is 0.5%).[7]

Theoretically, we might expect smaller countries to have relatively high own tax elasticities, as a change in their tax rate will not affect the market rate of return, making the cost of capital more responsive to tax changes (see Wilson (1999)). This may be one of the reasons that such countries are more likely to compete for corporate income using low rates; a change in the rate leads to a larger change in the relatively small tax base (see, for example, Bucovetsky and Haufler (2007)). The own tax elasticities in Table 4.2.2 show some evidence of this; they are higher for the Benelux countries than for France and Germany.

The importance of allowing for observed heterogeneity and correlation in locations' payoffs can be seen by looking at the cross tax elasticities. In a multinomial logit with no observed or unobserved heterogeneity all cross tax elasticities in a column would be the same - a reduction in the tax rate in location A would lead to patent applications switching from other locations in proportion to their original shares. This implausibly restrictive pattern of substitution is not imposed in our more flexible model, meaning that elasticities vary substantially within a column. In particular, our model allows the data to capture the fact that firms are more likely to choose to switch between locations with similar characteristics (whether that be because the firm has inventors located in several locations, or because locations have similar tax rates).

[7] Note that in Table 4.2.2 we adopt of the convention of reporting tax elasticities. Roughly speaking these can be interpreted as telling us the % change in the share of patents in location A associated with a 1% change in the rate of tax in location B. These can readily be converted to semi-elasticities (which, roughly speaking, tells us the % change in the share of patents in location A associated with a 1 percentage point change in the rate of tax in location B) by dividing by the appropriate tax rate. So for instance the 2005 rate of corporate tax in Germany was 38.3% and the German own tax elasticity is 0.201. Hence the German own tax semi-elasticity is given by 100x(0.201/38.3)=0.52%.

Table 4.5: *Own and cross price elasticities of locations with respect to statutory tax rates*

	Belgium	Denmark	Finland	France	Germany	Ireland	Italy	Luxembourg	Netherlands	Norway	Spain	Sweden	Switzerland	UK	US
Belgium	**-0.569**	0.010	0.018	0.085	0.174	0.001	0.016	0.006	0.079	0.003	0.003	0.030	0.036	0.025	0.038
Denmark	0.016	**-0.410**	0.015	0.044	0.145	0.000	0.013	0.007	0.060	0.003	0.004	0.033	0.047	0.025	0.037
Finland	0.015	0.008	**-0.465**	0.096	0.190	0.000	0.015	0.003	0.103	0.002	0.002	0.046	0.031	0.021	0.043
France	0.013	0.004	0.017	**-0.311**	0.118	0.000	0.009	0.002	0.052	0.001	0.001	0.017	0.017	0.011	0.021
Germany	0.007	0.004	0.009	0.032	**-0.201**	0.000	0.006	0.002	0.035	0.001	0.001	0.012	0.016	0.008	0.012
Ireland	0.048	0.031	0.033	0.103	0.279	**-0.404**	0.063	0.017	0.162	0.007	0.009	0.048	0.176	0.054	0.352
Italy	0.011	0.006	0.012	0.041	0.095	0.001	**-0.373**	0.004	0.043	0.002	0.002	0.017	0.035	0.019	0.017
Luxembourg	0.062	0.040	0.029	0.126	0.450	0.002	0.055	**-1.197**	0.125	0.012	0.019	0.058	0.093	0.073	0.079
Netherlands	0.022	0.010	0.034	0.094	0.234	0.001	0.017	0.004	**-0.569**	0.003	0.003	0.037	0.039	0.027	0.048
Norway	0.036	0.022	0.030	0.095	0.279	0.001	0.030	0.013	0.099	**-0.783**	0.009	0.082	0.070	0.046	0.050
Spain	0.012	0.009	0.008	0.032	0.097	0.000	0.012	0.008	0.047	0.003	**-0.336**	0.017	0.007	0.016	0.010
Sweden	0.017	0.012	0.032	0.066	0.166	0.000	0.015	0.003	0.078	0.004	0.002	**-0.432**	0.033	0.026	0.033
Switzerland	0.023	0.017	0.024	0.086	0.282	0.001	0.028	0.006	0.093	0.004	0.003	0.036	**-0.278**	0.029	0.049
UK	0.014	0.008	0.013	0.038	0.106	0.000	0.015	0.004	0.053	0.002	0.002	0.024	0.018	**-0.321**	0.030
US	0.013	0.008	0.018	0.048	0.102	0.002	0.009	0.003	0.059	0.002	0.001	0.020	0.025	0.019	**-0.439**

Notes: Each cell contains the elasticity of the share of patent applications in the country indicated in column 1 when the tax rate changes in the country in row 1. Numbers are calculated using tax rates and patent applications in 2005.

4.5 Policy simulations

One of the advantages of estimating the model outlined above is that it captures patterns of substitution across locations, and it therefore allows us to simulate counterfactual policy situations. We illustrate this by considering a recent set of policy reforms.

A number of European countries have introduced polices that offer substantially reduced rates of corporation tax on the income derived from patents, and in some cases other forms of intellectual property (these are often called Patent Boxes). Firms are able to declare that some portion of their profits are derived from either the use or licence of patents, and these profits are taxed at a lower rate. Patent Box rules differ across countries, for example, in terms of how eligible income is measured, how the rules that apply when calculating how much income can be allocated to patents, and how the related expenses are treated.[8] None of the countries require that the R&D underlying the intellectual property took place in that country, as this is not permissible under European law.

We use the most recent year of our data (2005) to simulate the impact of the two sets of policies. First we consider the introduction of Patent Boxes in the Benelux countries, and second the later introduction in the UK. We simulate the impact of these policies on the share of new patents for which legal ownership is placed in each of these countries using the choice model presented above. For illustrative purposes, we assume that the total level of patenting activity by European firms is not affected by the policy reforms. We also consider the impact of these policy reforms on tax revenue; this requires the further assumption that the relationship between where tax is levied and the location of legal ownership is not altered by the policy reform.

The policies are summarised in Table 4.6. In 2007 Belgium introduced a Patent Box that reduced the tax rate on income derived from patents from 34% to 6.8%, and the Netherlands introduced a Patent Box that reduced the rate from 31.5% to 10%. In 2008 Luxembourg reduced the rate from 30.4% to 5.9%. The UK government introduced a Patent Box at the rate of 10% in 2013; the main rate of corporate tax in the UK was 30% in 2005, but had fallen to 24% by 2013. We simulate the impact of the reduction from 30% to the Patent Box rate.[9]

[8]Evers et al. (2013) provide further details on the policies and incorporate the rules into effective tax rates.

[9]For further discussion of the UK Patent Box see Griffith and Miller (2010, 2011). A number of other European countries (Cyprus, Liechtenstein, Malta, Spain, and the Swiss canton of Nidwalden) have since introduced similar policies.

Table 4.6: *Patent Box regimes*

Applicant country	Year introduced (1)	Patent Box (2)	Effective rate (3)
Belgium	2007	Applies to gross income from patents and supplementary certificates	6.8%
Luxembourg	2008	Applies to net income from patents and some other forms of intellectual property	5.9%
Netherlands	2007	Applies to net income from patents and some other forms of intellectual property. Policy substantially broadened in 2010.	10%
United Kingdom	2013	Applies to net income derived from patents and similar types of intellectual property.	10%

Notes: Effective Patent Box rates are those that were in place when the policy was first introduced. Each policy is associated with criteria that defines which income is eligible. All policies include licence and embedded income. Policies differ in the conditions under which acquired intellectual property is eligible. Net and Gross refers to development costs. A number of other European countries now also operate policies akin to a Patent Box, and a similar policy has been proposed in the US.

Table 4.7 sets out the results of these simulations for the four locations that introduced Patent Boxes.[10]. A note of caution in interpreting these results is that the lowest tax rate we observe in the data is 10% in Ireland, whereas two of the Patent Box rates are below this level, and so are outside the observed range of taxes in our data.

We carry out the simulation on the full set of patent applications, shown in the top panel. It may be the case that many patents do not earn much income, and so in the bottom panel we carry out the simulation using only the high quality patents, under the assumption that these are the patents that are expected to earn the highest income. The estimates suggest that the location of these patents were on average more sensitive to tax.

The first column shows the actual share of patent applications in each location in 2005 (prior to the introduction of Patent Boxes). The second column shows the predicted share of patent applications in each location after the introduction of Benelux Patent Boxes. The standard error of these predicted shares are shown in parenthesis. The third column expresses the % change from column 1 to column 2. The introduction of Patent Boxes in the Benelux countries leads to a large and statistically significant increase in the share of new patents whose legal ownership is located in Belgium and Netherlands. The increase in Luxembourg is proportionally large, but is not statistically significant. There is no change in the share in the UK. There is a decline in the share of patent applications located in other non-Patent Box locations (not shown).

The fourth column shows the predicted shares after the introduction of the UK Patent Box (in addition to the Benelux Patent Boxes). The fifth column shows the % changes from column

[10]The full set of results, including the impact on other countries, is available on request

1 to column 4. The UK Patent Box leads to a reduction in the share of new patent applications made by subsidiaries located in the Benelux countries, but for Belgium and the Netherlands they still have a statistically significantly higher share than prior to the introduction of any Patent Boxes. The share of new patent applications made by subsidiaries located in the UK increases by a statistically and economically significant amount. The results with high quality patents are similar.

In columns (6)-(8) of Table 4.7 we consider the impact on tax revenue from income derived from patents. These combine two effects. The reduction in the statutory tax rate will reduce revenue, but the increase in the share of income from patents will increase it. We demonstrate the impact on tax revenue by computing the product of the statutory tax rate in each country and the share of patent applications. We index this to 100 before the introduction of any Patent Boxes. In the upper panel of the Table we assume that all patents are equally valuable, and that the relationship between legal ownership and taxable income is not affected by the reform. All countries experience a decline in revenue. Although the countries that introduce Patent Boxes attract more new patents, the increased share is not sufficient to outweigh the effect of the lower tax rate. With all four Patent Box policies in place, revenues are less than half of their previous levels in these countries. Ernst and Spengel (2011) provide evidence that lower rates of tax on patent income attract particularly innovative projects with high earning potential. In the lower panel we consider the effect on revenue when we consider only high quality patents. The picture is here is similar; the introduction of Patent Boxes results in a substantial reduction in revenues.

Table 4.7: *Impact of Patent Boxes on location of and tax revenue raised from new patents*

	Share of new patent applications					Tax revenue		
	Pre reform	Benelux Patent Boxes	% change	UK Patent Box	% change	Pre reform	Benelux Patent Boxes	UK Patent Box
All patents								
Belgium	2.39	3.53	47.6%	3.42	-3.1%	100	30	29
		(0.34)		(0.34)				
Luxembourg	0.33	0.60	83.9%	0.56	-7.0%	100	36	33
		(0.34)		(0.34)				
Netherlands	7.92	12.51	58.0%	12.16	-2.8%	100	50	49
		(0.38)		(0.38)				
UK	4.15	4.15	-0.1%	5.25	26.6%	100	100	42
		(0.31)		(0.32)				
High quality patents								
Belgium	1.90	3.29	73.2%	3.17	-3.9%	100	35	33
		(0.35)		(0.35)				
Luxembourg	0.42	0.76	82.8%	0.71	-7.3%	100	35	33
		(0.35)		(0.35)				
Netherlands	7.00	12.48	78.3%	12.09	-3.1%	100	57	55
		(0.39)		(0.38)				
UK	4.89	4.39	-10.2%	5.64	28.5%	100	90	38
		(0.31)		(0.33)				

Notes: The top panel provides numbers based on all patent applications; the bottom panel provides numbers on high quality patents only. Column 1 shows the actual share of patent applications in each location; columns 2 and 4 give the predicted share of patent applications in each location following the introduction of the Benelux Patent Boxes and following the additional introduction of the UK Patent Box. Standard errors are in parenthesis. Column 3 and 5 show the corresponding percent changes in shares relative to column 1. The final three columns show revenue raised from new patents, assuming that all patents have equal expected values. Revenues are indexed to 100 in the pre Patent Box period. Numbers are based on simulations using data for 2005.

4.6 Summary and conclusion

The literature has emphasized the downward pressure on corporate income tax rates that arises from factor mobility. There is also a large literature that discusses the strategies firms use to shift income for tax purposes and to circumvent anti-avoidance rules, and that highlights an important role for intangible assets. However, we know relatively little about the extent to which the location of intangible assets responds to tax. The evidence there is on the impact of tax on the location of capital more generally has tended to suffer from the imposition of restrictive a priori assumptions placed on the underlying model of firm behavior. From a policy perspective it is clearly important to understand how responsive firms are to corporate income taxes when they make location decisions.

In this paper, we estimate a model of firms' decisions over where to locate the legal ownership of their patents. We find that corporate tax rates are an important determinant of location choice. We extend the current literature on the determinants of firm location choice by estimating a flexible choice model, which accounts for both observed and unobserved heterogeneity in behavior. We are able to generate own and cross tax elasticities across locations that capture complex patterns of substitution in the data. The model can be used to conduct *ex ante* analysis of policy changes. We find that this heterogeneity is important for explaining location choices.

Our model also shows that other factors influence where firms choose to hold legal ownership of patents. For instance, firms are more likely to locate patent ownership in countries where they have associated real innovative activity. This may reflect co-location externalities, or the influence of tax rules which seek to limit the extent to which income and real innovative activity can be geographically separated. Firms also value other non-tax location characteristics. Such factors, along with tax rules like the operation of CFC regimes that limit the tax advantages of locating patent ownership in low tax jurisdictions, help explain why we do not see firms choosing to hold all legal ownership of patents in the lowest tax locations.

We use the model to consider the impact of the recent introduction of preferential tax regimes for income from patents. These Patent Boxes are likely to attract patent income, but our estimates suggest they will also lead to substantial falls in tax revenues. Of course some of this revenue loss might be offset by gains from attracting activities that yielded positive externalities; these would need to be taken into account in a calculation of the welfare impact of the policy. It is also possible that the tax reforms will affect firms decisions over whether to apply for a patent on a new technology or whether to rely on secrecy. We do not have information that would allows us to directly estimate this margin, but this would be an interesting avenue for future research.

The introduction of Patent Boxes by several European countries in a relatively short space of time has given rise to concerns that countries are engaging in tax competition for patent income. In future work we intend to build on the framework developed here to consider whether governments are engaged in a strategic game to attract income from intellectual policy that ultimately will continue to exert downward pressure on corporate taxes.

Chapter 5

Partial consumption smoothing over the Great Recession

5.1 Introduction

Over the Great Recession households in the US and UK experienced adverse shocks to income and large increases in the price of food. Unlike previous recessions, there was a substantial fall in real expenditure on food, which has led some to infer a substantial reduction in the size and nutritional quality of households' food baskets (see, for example, Taylor-Robinson et al. (2013) for concern about the rising rates of food poverty in the UK and US Department of Agriculture (2013) for the US and US Department of Agriculture (2010) and Lock et al. (2009) for concerns that households are buying cheaper, less nutritious calories). However, it is well known that equating expenditure with consumption can lead to mistaken conclusions about how households are affected by changes in their economic environment (e.g. Aguiar and Hurst (2005)). The ability of households to insure themselves against income shocks is a question of central concern in economics (Blundell et al. (2008), Blundell and Preston (1998), Jappelli and Pistaferri (2010), Hall and Mishkin (1982), amongst others). Shocks to wages alter the opportunity cost of time, in which case households may switch away from market goods towards greater time spent in home production (Becker (1965)), or they may allocate more time to searching out lower prices for a fixed basket of goods (Stigler (1961)). In addition, households may be able to substitute away from some characteristics in order to smooth consumption of others; for example, households may switch from a preferred branded to a cheaper generic product in order to maintain the nutritional quality of their food basket. We are interested in the extent to which households are able to exploit these mechanisms to smooth, or "insure", the quantity and nutritional quality of their food basket in the face of adverse shocks.

Our contribution to the literature is twofold. First, we show that households were able to maintain the number of calories and their nutritional quality over the Great Recession by acting to reduce the (real) price that they paid for their shopping baskets. We do this using detailed household level transaction data from the UK. Second, we set out a model of grocery shopping behavior to help us understand the mechanisms that households used to do this. This extends Aguiar and Hurst (2007), who show in a cross-section that observed reductions in expenditure at retirement do not necessarily equate to a reduction in consumption, but rather, as an individual's opportunity cost of time declines at retirement they switch away from market goods and towards home production and increased search. We build on this approach by also incorporating the possibility that households can adjust the characteristics of their shopping basket to lower the price of the basket. We show that households were able to smooth two aspects of consumption: the nutritional quantity (calories) and nutritional quality of their food purchases. They achieved this by using time (to search out better deals), by switching away from their preferred non-nutritional characteristics (for example, from branded to generic products) and by substituting away from more expensive foods and nutrients (such as alcohol and protein) towards cheaper ones. These adjustments meant that, although households were made worse off as a result of the recession, the nutritional quality of their food purchases did not decline. There is evidence from the US that as economic conditions worsen households spend longer shopping and pay lower prices (Kaplan and Menzio (2014b)), increase their use of sales, switch to generic products (Nevo and Wong (2014)) and switch to low-price retailers (Coibion et al. (2014)). We also show that UK households adjusted the nutritional composition of their shopping baskets and they did this in such a way as to maintain (and in fact slightly improve) its overall nutritional quality.

Our work relates to several literatures. Most closely related is a series of influential papers by Aguiar and Hurst (2005, 2007) who take a similar approach applied to a different setting. They consider household behavior around the time of retirement. Bernheim et al. (2001) argue that the observed decline in expenditure at retirement is evidence that households do not plan adequately for retirement, and by implication suffer a fall in living standards. Aguiar and Hurst use data on food purchases to show that, by increasing their time spent searching and on home production, households are able to maintain a similar level of food consumption in retirement, despite spending less. We extend their analysis to consider how households respond to unpredictable changes in incomes and prices by substituting across basket characteristics, as well as by increasing their time spent searching for lower prices. Also related is Aguiar et al. (2013), who show with time use data on US households that over the Great Recession 30%

of foregone market work hours were allocated to non-market work, and 7% were allocated to increased shopping effort. We relate our findings to theirs by using our model to infer the opportunity cost of time and show that it fell over the Great Recession. In a recent extension to this literature, Nevo and Wong (2014) show that US households increased their time spent shopping and in home production, so that the decline in consumption was substantially less than the decline in food expenditure.

Also related to this paper is the literature on insurance and consumption smoothing in an intertemporal setting. These papers typically focus on the response of consumption to permanent and transitory shocks to income (see, Blundell et al. (2008), Blundell and Preston (1998), Jappelli and Pistaferri (2010), Hall and Mishkin (1982), among others). This body of work studies how households can transfer income intertemporally to smooth consumption. However, Blundell et al. (2014) show the importance of family labor supply as an insurance mechanism to wage shocks; once this, and taxes are properly accounted for, there is little evidence of additional insurance. They consider a lifecycle setup in which households choose consumption and leisure to maximize their utility; the optimal choices made by households are such that consumption is smoothed following wage shocks. We are interested in understanding the smoothness of two aspects of consumption – the nutritional quantity and quality of households' shopping baskets – and how this can result from the *intra-temporal* utility maximization of households. We show that the ability of households to re-optimize over the quantity of food, its characteristics and the time spent shopping is crucial for understanding consumption smoothing over this period.

Our results contribute to those found in the literature which suggest that nutrition and health might improve as economic conditions worsen. Strauss and Thomas (1998) show that the effect of economic shocks on nutritional status (energy intake, weight, child stature) in Russia in the late 1990s were such that individuals and households were, *"able to weather short-term fluctuations in economic resources, at least in terms of maintaining body mass index and energy intake,"* and that individuals switched to cheaper and less tasty calories in hard times. By studying variation over time across US states, Ruhm (2000) shows that diets become less healthy and obesity increases when the economic situation improves. Dehejia and Lleras-Muney (2004) find that babies conceived in recessions have a lower probability of bad outcomes such as low birth weight, congenital malformations, and post-neonatal mortality. However, Adda et al. (2009) show that permanent income shocks have little effect on a range of health outcomes.

We begin in Section 5.2 by describing our data and showing how expenditure, calories and nutritional quality evolved over the Great Recession. In Section 5.3 we outline a simple optimizing model of consumer grocery shopping and set out our empirical strategy. Section 5.4 describes how we measure households' choices of shopping effort and basket characteristics. Section 5.5

presents empirical estimates of the price function and quantitative estimates of how households were able to maintain calorie purchases in the face of lower real food expenditure. A final section summarizes and concludes.

5.2 Food expenditure and consumption

We use information on food (including drinks and alcohol) that is purchased and brought into the home by a representative panel of British households over the period January 2005–June 2012. The data are from the Kantar Worldpanel and are collected via in-home scanning technology. Participants record spending on all grocery purchases via an electronic hand held scanner in the home. Purchases from all types of store – supermarkets, corner stores, online, local speciality shops – are covered by the data. The data include information on the exact price paid for the product, whether or not the product purchased was on promotion (e.g. ticket price reduction, "Buy One Get One Free", etc.), nutritional information (number of calories, amount of salt, protein, saturated fat and other information that is listed on food labels) and demographic details of the households. These data have been used in Dubois et al. (2014) and Griffith et al. (2009), and similar data are widely used in the US, for example in Aguiar and Hurst (2007); see Griffith and O'Connell (2009) and Leicester and Oldfield (2009) for further discussion of the data. Our sample includes 14,694 households and over 450,000 "shopping baskets", which we define as all purchases made by a household in a month.

5.2.1 Real food expenditure and calories

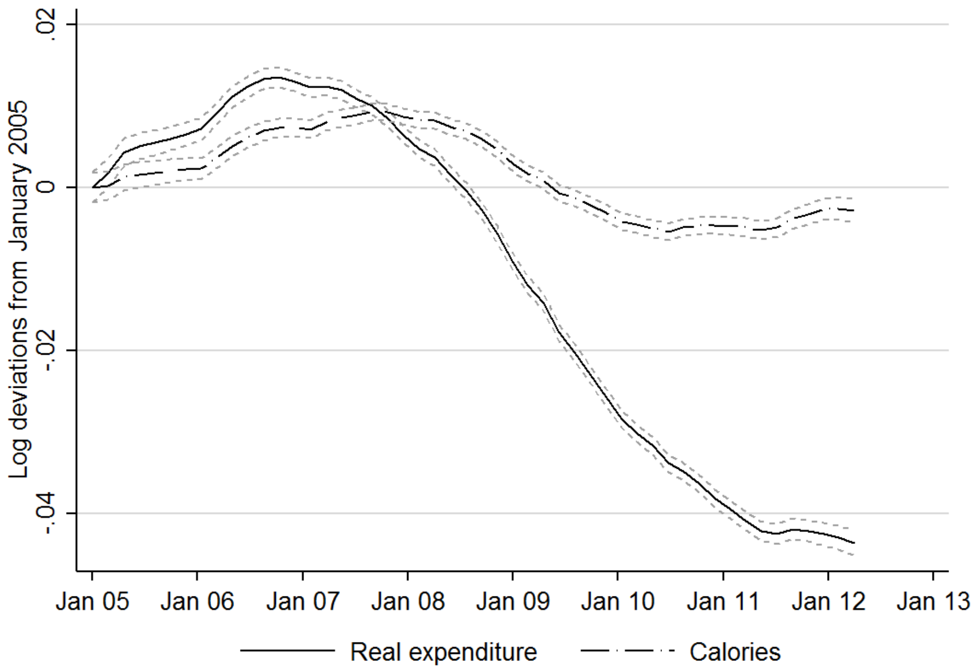
Our focus is on the grocery baskets that households purchase for home consumption, which constitute over 85% of total calories purchased (see Section 5.5.3). Figure 5.1 shows how real food expenditure and calories changed over 2005-2012. Real food expenditure is nominal expenditure on food and drink deflated by the food and drinks component of the Consumer Price Index (CPI). Calories are expressed per "adult equivalent" per day and real expenditure per "adult equivalent" per month.[1] The figure shows log deviations from the first quarter in our data (2005Q1) based on with household variation; it shows that there was a sharp decline in real food expenditure in 2008. Households reduced real grocery expenditure by over 6% between the pre-recession years, 2005-2007 and the period post-recession, 2010-2012. However, calorie purchases remained reasonably smooth over this period (falling by only 1% between 2005-2007

[1] As in Dubois et al. (2014) we "equivalize" to account for differences in household size and composition using an "adult-equivalent index" based on the estimated average requirement (EAR) for energy of household members (Department of Health (1991)), which vary by age and sex. We sum the EARs of all household members and divide by 2550; this equals 1 for a household containing only one adult male aged 19-59. If the household contained one adult male, one adult female (EAR=1940) and one female infant (EAR=698) then the index would be 2.035=(2550+1940+698)/2550; this means that if the household purchased 5188 calories this would be "equivalized" to 2550 and so be comparable to a single adult male purchasing 2550 calories.

and 2010-2012, see also Table 5.1). The fact that households reduced calories by less than their real food expenditure indicates that they switched toward cheaper (in real terms) calories. The focus of this paper is on how they achieved this price reduction.

Aggregate food consumption in the US behaved in a similar way to the UK. According to US Department of Agriculture (2014) per capita food expenditure, in constant prices, fell by 4% between 2007 and 2009, and there is some evidence that consumers maintained a similar quantity of food, but switched to cheaper, less convenient alternatives and private label brands (see Kumcu and Kaufman (2011) and Kuchler (2011)).

Figure 5.1: *Real food expenditure and calories purchased*



Notes: The figure shows the log deviations in real expenditure and calories relative to 2005Q1. Numbers are based on within household variation. Real expenditure is nominal expenditure on food at home deflated by the food and drink component of the CPI in 2008 prices. Numbers are expressed per adult equivalent. Lines are local polynomials with 95% confidence intervals shown as dotted lines.

Over the Great Recession households experienced different shocks. For example, Crossley et al. (2013) show that younger households were particularly hard hit. In the UK, the incomes of households towards the bottom of the income distribution were largely protected from the immediate impact of the Great Recession by the benefit system (Brewer et al. (2013)). It is possible that the smoothness in calories seen at the average masks differences across households. We look at the changes in real expenditure and calories purchased by demographic composition of the household and by the employment status and income of the household.

We distinguish households by whether they include pre-school children, school-aged children (and none at pre-school ages), adults (non-pensioner households without children), and pensioner

Table 5.1: *Changes in real food expenditure and calories, per adult equivalent*

	Real expenditure (£ per month)			Calories purchased (per day)		
Households	2005-2007	2010-2012	% change	2005-2007	2010-2012	% change
All	114.52	107.27	-6.33	2300	2274	-1.10
pre-school children	94.15	82.21	-12.68	2011	1931	-3.99
school aged children	93.00	83.60	-10.10	2041	1948	-4.57
adults	116.65	110.72	-5.08	2288	2295	0.29
pensioners	129.09	121.69	-5.73	2530	2497	-1.32
working high income	111.43	102.68	-7.85	2028	2011	-0.86
working mid income	108.41	99.72	-8.02	2150	2099	-2.37
working low income	98.97	92.51	-6.53	2170	2131	-1.81
unemployed	105.64	98.70	-6.57	2271	2230	-1.78

Notes: Real expenditure is nominal expenditure on food at home deflated by the food and drink component of the CPI in 2008 prices. Real expenditure is per adult equivalent per month; calories are per adult equivalent per day. % changes refer to the average within household percentage change. "Pre-school" denotes households with a child aged between 0 and 5; "school age" are households with the youngest child between 6 and 17. "Adults" are households where everyone is 18 or older and everyone is aged below 65. "Pensioner" households are those in which at least one member is aged 65 or over. Working households are those in which the head of the household works more than 8 hours a week. Income is measured using information on occupation and education contained in social grade; grade AB/C/DE correspond to high/middle/low income. The percentage change is the average within-household change in each variable.

households. There is considerable policy interest in how households with young children have been affected by the recession. For example, US Department of Agriculture (2013) argue that in the US food insecurity is more prevalent in households with children under six than in the whole population, and changes in food purchasing decisions, particularly those that affect nutritional quality, may have important health consequences for young children (see, for instance, Currie (2009) and Case et al. (2005)).

Table 5.1 shows the levels in 2005-2007 and 2010-2012 and percentage changes in real monthly expenditure on food at home per adult equivalent and calories purchased per adult equivalent per day for the different household types.[2] On average, the nominal food expenditure of all household types failed to keep pace with the rise in food prices, meaning that real expenditure fell. Households with pre-school children reduced real expenditure by the most at 12.7%; households with school age children also experienced a relatively large reduction of 10.1%. In addition, households with children (both pre-school and school age) reduced the number of calories that they purchased per adult equivalent, although by much less than real expenditure. This is in contrast to households without children, who reduced real expenditure by about half the amount as households with children. Adult households did not, on average, reduce calories, while pensioner households reduced calories by less than one-third the amount that households

[2]For reasons of parsimony, in tables throughout the paper we compare the period 2005-2007 with 2010-2012. The intervening period, 2008-2009, was characterized by reductions in real incomes and rising food prices; after 2009 incomes remained depressed and the food price level remained high. Typically numbers for 2008-2009 lie somewhere in between numbers for the pre- and post-recession periods.

with children did.[3] Despite differences in the magnitude of the changes across households, smoothing is evident for all household types: calorie purchases declined by much less than the falls in real expenditure.

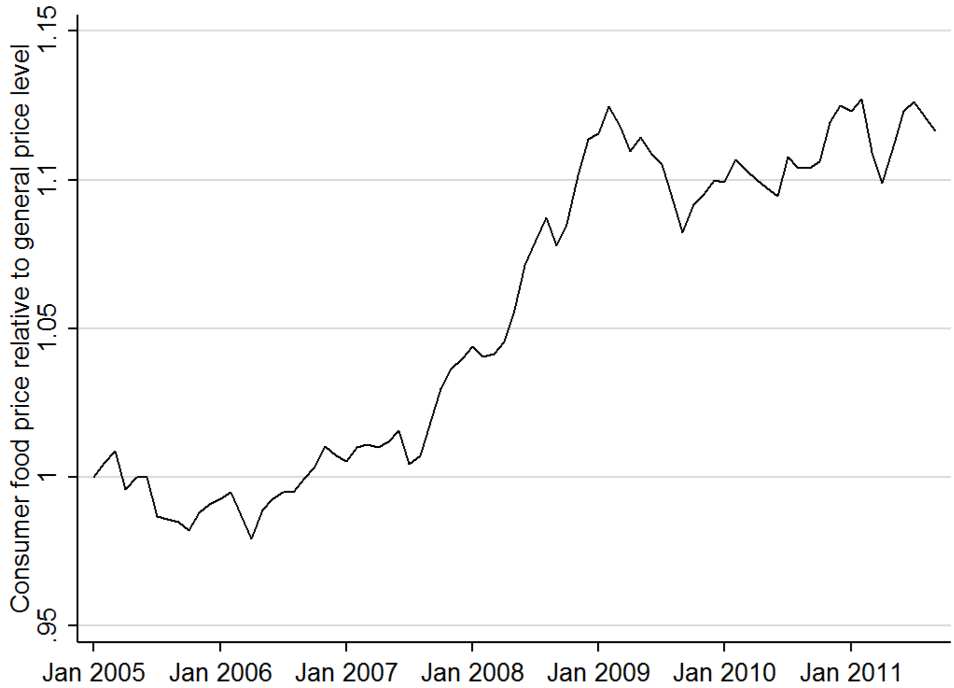
We also look at how these patterns vary by the income level of the household. We group households according to their work status and income: households in which the head of the household works more than 8 hours a week are deemed to be "working", the remaining households are either "unemployed" or pensioner households. "Working" households are further divided by income. We use information on the occupation and eduction of the main earner contained in a variable called social grade to measure income.[4] High income households include higher and intermediate managerial, administrative and professional occupations (social grades A and B); middle income includes clerical and junior managerial, administrative and skilled manual occupations (C), and low income include semi- and unskilled manual workers (D and E). There is a strong correlation between income and the social grade classification – on average, households in social grade A have a main income earner with a net annual income of almost £40,000, whereas those in grade E have a main income earner with a net annual income of less than £5,000.

Reductions in real food expenditure are largest for working households with high and middle levels of income. However, working households with higher levels of income cut back on their calories by the least, while working households at the middle of the income distribution reduced their calorie purchases by the most, indicating that high income working households reduced the price per calorie they paid for their groceries by more. The numbers shown are evidence for smoothing of calorie purchases by households across the income distribution – the real expenditure of all groups declined, but calorie purchases fell by much less.

[3] One potential concern is that, because we are looking *within* household, as children age they may purchase more foods outside of the home, and this might in part be driving our results. To check this we use repeated cross-sectional data from the *Living Costs and Food Survey 2005-2011* and find that the change in total calories (from all food) per adult equivalent per day is -2.9% for households with pre-school children and -3.5% for households with school age children.

[4] See http://www.nrs.co.uk/nrs-print/lifestyle-and-classification-data/social-grade for details.

Figure 5.2: *Real consumer price of food*



Notes: The figure shows the Consumer Price Index for food relative to the Consumer Price Index for all items over 2005-2011.

The data suggest that the experience of households of all types was similar – large declines in real food expenditure were accompanied by much smaller falls in calorie purchases. This is because, although different households experienced different income (wage and asset price) shocks, all households were subject to higher food prices – from 2005-2007 to 2010-12, the consumer price of food rose by 10% more than the consumer price of all goods (see Figure 5.2). It is likely that this price shock was the main reason that households' real food expenditure fell. The stability of calorie purchases over this period is due to households switching to cheaper (in real terms) calories. In Section 5.5 we investigate the mechanisms by which they did this; however first we describe how the nutritional quality of households' grocery baskets changed over this period.

5.2.2 Nutritional quality

In the previous section we showed that, although real expenditure declined markedly over the recession, the number of calories that households purchased remained relatively stable. Households achieved this by lowering the average real price per calorie that they paid for their shopping basket. A possible concern is that a switch to cheaper calories could lead to a reduction in the nutritional quality of those calories (see, *inter alia*, Lock et al. (2009) and US Department of Agriculture (2010)). It has been well documented (e.g. US Department of Agriculture (1997), US Department of Agriculture (2000)) that there are cross-sectional differences in the nutritional quality of food purchases, with richer households purchasing food of a higher nutritional quality, on average. We observe this in our data, but our focus is to consider the *within*-household variation in the nutritional quality of food purchased over the Great Recession.

Measuring nutritional quality is complex; households made changes that improved nutritional quality in some dimensions and reduced nutritional quality in other dimensions. For example, over the recessionary period, the share of calories from protein fell for almost all households; this is generally considered to be "bad" for nutritional quality, as most UK households purchase less protein than the recommended amounts. In the other direction the share of calories from saturated fat declined; this is generally considered to be "good" for nutritional quality, because most households purchase more saturated fat than the recommended amounts. These changes in the nutritional composition of shopping baskets are such that it is not immediately obvious whether nutritional quality improved or worsened over this period.

To gain a better understanding of the overall changes in the nutritional quality of households' shopping baskets we use the United States Department of Agriculture's (USDA) Healthy Eating Index (HEI) (see US Department of Agriculture (2007)). The HEI gives a score between 0 and 100 based on the density (i.e. amount per 1000 calories) of different food groups and nutrients in a basket. US Department of Agriculture (2007) comment that density standards are appealing, *"not only because they allow a common standard to be used, but because they have the advantage of being independent of an individual's energy requirement."* This means that changes in the HEI will largely abstract from changes in the quantities of nutritional components purchased that arise due to changes in the total number of calories purchased. The HEI is used by Beatty et al. (2014) to analyze changes in the dietary quality of the US population over the 1989-2008 period.

We are interested in how the nutritional quality of a household's shopping basket compares to the one they purchased prior to the Great Recession. We calculate the average within household change in the HEI and its component scores between 2005-2007 and 2010-2012, shown in Table 5.2. The overall average HEI increases by around 1.5% over this period, though this is small relative to the cross sectional variation; the standard deviation of the HEI across households is 10. However, it represents an aggregation of some larger changes that go in offsetting directions, for example, a shift away from vegetables, grains, milk and meat was offset by a reduction in the saltiness of food purchased and a lower calorie share of saturated fat. This suggests that, although households adjusted the relative composition of nutrients and food groups in their baskets, potentially in ways that reduced their utility, they did so in such a way as to maintain the average level of nutritional quality in the basket.

Table 5.2: *Changes in the Healthy Eating Index*

	Max score	Mean in 2005-2007	Change to 2010-2012
HEI 2005-2007	100	49.0	0.72
of which			
"Good" change			1.45
"Bad" change			-0.72
which consists of:			
Total fruit	5	3.06	-0.02
Whole fruit	5	3.36	0.08
Total vegetables	5	3.20	-0.13
Dark green/orange veg	5	1.61	0.00
Total grains	5	3.69	-0.03
Whole grains	5	1.55	-0.11
Milk	10	5.28	-0.05
Meat	10	7.96	-0.22
Oils	10	4.93	-0.18
Sodium	10	6.42	0.93
Saturated fat	10	2.70	0.27
Calories from SoFAAS	20	5.22	0.18

Notes: Column 1 shows the maximum score for the overall HEI and each component; column 2 shows the mean of the overall HEI and the component scores in 2005-2007; column 3 shows the mean within household in the scores to 2010-2012. "Good change" (shown in row 2) is the sum of the positive changes in the bottom panel; "Bad change" (shown in row 3) is the sum of the negative changes in the bottom panel. "Calories from SoFAAS" is the share of calories from solid fat, added sugar and alcohol.

Table 5.3: *Changes in the Healthy Eating Index, by type of household*

Households	Mean HEI 2005-2007	Change to 2010-2012	*of which:* (%)	"Good"	"Bad"
All	49.0	0.72	(1.5%)	1.45	-0.72
pre-school children	48.7	1.52	(3.1%)	3.02	-1.51
school aged children	46.1	1.03	(2.2%)	1.90	-0.87
adults	47.8	1.46	(3.1%)	1.93	-0.46
pensioners	51.5	-0.23	(-0.4%)	0.91	-1.14
working high income	49.6	0.87	(1.8%)	1.78	-0.91
working mid income	48.0	1.03	(2.1%)	1.78	-0.75
working low income	46.6	2.01	(4.3%)	2.44	-0.43
unemployed	46.7	1.11	(2.4%)	1.67	-0.56

Notes: Column 1 shows the mean HEI score for each households group in 2005-2007; the second column shows the mean within household change to 2010-2012 within each group; column 3 shows this in percentage terms. Columns 4 and 5 show the "Good" and "Bad" changes calculated within each group in the way described in Table 5.2. Household group definitions given in the notes to Table 5.1.

We also look by the different household types, see Table 5.3 (a more detailed breakdown is shown in Table C.1 in the Appendix Section C.1). Households with pre-school children improved their HEI score by the most: despite a relatively large fall in the contribution of vegetables and meat, they improved with respect to fruit, salt, saturated fat and alcohol by more than enough to compensate. The HEI score of the shopping baskets of pensioner households declined slightly: unlike households with pre-school children, they did not decrease their saturated fat purchases by enough to compensate for the switch away from meat and vegetables. However, pensioner households had the highest HEI scores to begin with. There is a cross-sectional correlation between average nutritional quality and income: households with higher incomes have a higher HEI score than households in the lowest income band. The magnitude of the difference in the average HEI score between high and low income households is similar to that found by Beatty et al. (2014) in the US. However, low income working households improved the nutritional quality of their shopping basket by more than working households with higher income; primarily by switching towards fruit, away from saturated fat and alcohol and reducing the salt content of their grocery purchases.

Overall, it seems that households were not only able to smooth the number of calories that they purchased, but also maintain the nutritional quality of these calories. In the next section we set out a model of grocery shopping in which households choose the number of calories, the characteristics of these calories and their shopping effort to maximize their utility. We use this model to determine how households were able to adjust their consumption and shopping effort to maintain the number and nutritional quality of calories that they purchased.

5.3 A model of grocery shopping

5.3.1 Model

We model the decisions that a household makes over its grocery shopping. Our set up shares a number of features in common with that in Aguiar and Hurst (2007); households choose the total amount of groceries to buy and how much time to allocate to shopping and home production (specifically, cooking). Spending more time shopping allows households to lower their expenditure on groceries, but they incur a cost of time. We extend Aguiar and Hurst (2007) to also model the choice a household makes over the characteristics of their grocery basket. We are particularly interested in its nutritional characteristics. This modification turns out to be important for studying how households adjust their shopping behavior in response to economic shocks.

We model the household's utility from food consumption (v) as depending on the total number of calories in its shopping basket, C, and a K dimension vector of basket characteristics, $\mathbf{z}$. Grocery basket characteristics include the nutritional and food group composition of the basket, the share of the basket from branded products, and the time required to prepare calories for consumption (we denote this by z', which is an element of $\mathbf{z}$). Note that inclusion of calories in the objective function does not imply that relaxation of the consumer's budget constraint will translate directly into more calories. Calories is one argument of many in the consumer's utility function – the consumer will trade off a larger shopping basket with improvements in the nutrient and quality content of the basket. In addition, the relationship between utility and calories, all else equal, may be highly concave – at low level more calories may increase utility by a large amount, at moderate or high levels more calories may increase utility only infinitesimally.

We denote the price that the household pays per calorie for its grocery basket $P = P(e, \mathbf{z}; \phi)$. P depends on how much effort the household expends shopping, e. All else equal, more time shopping results in a lower price paid for groceries, because the shopper finds better deals (that is we expect $\partial P/\partial e < 0$, although it is likely that there are diminishing returns to shopping effort, meaning $\partial^2 P/\partial e^2 > 0$). The characteristics of the shopping basket, $\mathbf{z}$, can also affect the price paid per calorie. For example, increasing the share of calories from protein will likely increase the price per calorie, while increasing the share of generic rather than branded products will likely decrease the price per calorie. Finally, we denote by ϕ other factors that affect the price per calorie the household pays for its groceries, including for example, common time varying factors, such as the prices at which firms offer food in the market, regional-time varying factors, such as local market conditions, household level characteristics, such as shopping efficiency, and household-time varying characteristics, such as caloric requirements of the household.

Spending more time shopping has the advantage of potentially lowering the household's monetary expenditure on groceries, but it has the downside of leaving less time for the household to engage in leisure or market work. We denote the opportunity cost of time by ω. Like other characteristics of the grocery basket, the preparation requirement may affect the price per calorie, but unlike other characteristics preparation is also costly in terms of time.

We assume that preferences over total calories and characteristics are weakly separable from other arguments in the household's utility function, and that choices other than those over $(e, \mathbf{z})$ do not enter directly into the price function. This implies that changes in work status affect household's choices through changing the resources that are available to spend on food and the opportunity cost of time, but not through altering the relative desirability of different basket characteristics or the marginal rate of substitution between calories and any given characteristic.

The household's problem can be stated as a cost minimization problem given by:

$$\min_{e,\mathbf{z},C} P(e, \mathbf{z}; \phi)C + \omega(e + z'), \tag{5.1}$$

$$s.t. \quad v(C, \mathbf{z}) = \bar{v}. \tag{5.2}$$

The household's choice over consumption of non-food and over leisure and labor supply are captured in the opportunity cost of time ω, and the total resources allocated to food consumption is captured in $\bar{v}$. We assume that the household does not select zero shopping effort ($\partial p/\partial e \rightarrow$

$-\infty$ as $e \to 0$ ensures this), or zero leisure or cooking time (appropriate Inada conditions on the utility function ensure this).

The first order condition for shopping effort is:

$$-\frac{\partial P}{\partial e} C = \omega, \tag{5.3}$$

i.e. the household puts effort into shopping up to the point where the marginal gain in terms of lower food expenditure equals the opportunity cost of time. This optimality condition can be used to infer the household's opportunity cost of time, providing a measure that has the advantage that it allows us to remain agnostic about the workings of the labor market. The first order condition for the choice of total calories is:

$$P = \lambda \frac{\partial v}{\partial C}, \tag{5.4}$$

where λ is the Lagrange multiplier on the household's constraint (5.2) and can be interpreted as the reciprocal of the marginal utility of more resources allocated to food consumption (either an extra £ of expenditure or an extra £ worth of time spent shopping). Condition (5.4) says that the household will select the number of calories that equates the marginal cost of more calories with the marginal utility of calories (converted into monetary terms through multiplication by λ).

The first order condition for the choice of characteristic k (where $z_k \neq z'$) is:

$$\frac{\partial P}{\partial z_k} C = \lambda \frac{\partial v}{\partial z_k}. \tag{5.5}$$

Interpretation is similar to the calorie first order condition: for each characteristic k, the household will choose the quantity that equates its marginal cost with the marginal utility from that characteristic (expressed in monetary terms). For the cooking requirement characteristic, the first order condition is $\left(\frac{\partial P}{\partial z'} C + \omega\right) = \lambda \frac{\partial v}{\partial z'}$.

The ratio of condition (5.5) and (5.4) yields the marginal rate of substitution between calories and characteristic k:

$$\frac{\partial v / \partial z_k}{\partial v / \partial C} = \frac{\partial P}{\partial z_k} \frac{C}{P}. \tag{5.6}$$

At the optimum, the number of extra calories the household needs as compensation for a marginal loss in the amount of characteristic k to remain indifferent to the change equals the ratio of the marginal costs of characteristic k and calories.

This framework is well-suited to studying how households adjust their shopping behavior in response to deteriorations in the economic environment that they face. We use the model to analyse changes over the period spanning the Great Recession. Households in the UK experienced reductions in their real incomes, driven by slow nominal wage growth and reductions in asset prices; in the US there were also substantial falls in real incomes, although rising unemployment played a more central role. Importantly, households also faced much higher food prices. In problem (5.1)-(5.2) this would lead to changes in the resources the household had available for food consumption, v, the opportunity cost of time, ω and the market prices of foods, captured by ϕ.

The negative economic shocks experienced over the recession led to a reduction in $\bar{v}$, meaning that households were made worse off. However, we observe empirically that the number of calories purchased by households and the nutritional quality of these calories remained stable. We are interested in how households were able to adjust their time spent shopping and *other aspects* of consumption, e.g. the share of their calories from generic products, in order to smooth the size and nutritional quality of their shopping baskets. How households can do this can be illustrated by a simple example. Suppose that a household gets utility from a good that is branded, z_b, a generic good, z_g, and a nutrient characteristic, z_n, that is provided in differing degrees by each good. Following an inward shift of their budget constraint the household shifts to a lower indifference curve but will also adjust the relative consumption of z_b and z_g; possibly adjusting z_b and z_g to maintain z_n (analogous to the number of calories, or their nutritional quality), despite being made worse off.

Our empirical strategy is to specify a parametric form for the price per calorie function $P(e, \mathbf{z}; \phi)$ and use this to estimate the sensitivity of the price per calorie that households paid for their grocery baskets to the choice variables $(e, \mathbf{z})$.

5.3.2 Empirical functional form

At this point it is useful to introduce a household index h and a time index t. We have panel data on households' daily food purchases, but to consider the household's entire shopping basket we aggregate each individual household's purchases to the monthly level; we observe each household for many months (on average 31 months). We measure the price per calorie that household h pays for its groceries in period t, P_{ht}, as a weighted average of the transaction prices that the household pays for the individual products in its grocery basket. Let i index a product (i.e. a barcode or UPC), s index a store and d index a date. Let c_i denote the number of calories in product i and p_{isd} the market price of product i in store s on date d. P_{ht} is given by:

$$P_{ht} = \sum_{isd \in t} \left(\frac{p_{isd}}{c_i} \right) w_{hisd}, \tag{5.7}$$

where $\frac{p_{isd}}{c_i}$ is the price per calorie of product i in store s on date d. The weights are given by:

$$w_{hisd} = \frac{c_i b_{hisd}}{\sum_{i's'd' \in t} c_{i'} b_{hi's'd'}}, \tag{5.8}$$

where $b_{hisd} \in \{0, 1, 2, ...\}$ is the number of purchases of product i from store s on date d by household h. It is through their choice of products, b_{hisd}, that households are able to change the average price they pay per calorie. Similarly, each characteristic of the shopping basket is defined as a weighted average of the "amount" of the characteristic in each product in the basket.

Total calories purchased by a household in a month is given by:

$$C_{ht} = \sum_{isd \in t} c_i b_{hisd}. \tag{5.9}$$

We do not directly observe the time that a household spends shopping; we use a vector of shopping trip characteristics to proxy shopping effort, outlined in Section 5.4.1.

As our baseline specification we assume that the price function, $P(\mathbf{e}, \mathbf{z}; \phi)$, can be approximated by a log-log specification (Triplett (2004), Aguiar and Hurst (2007)); we show in the robustness section that our results are robust to an alternative polynomial specification. Specifically, we consider:

$$\ln P_{ht} = \alpha \ln \mathbf{e_{ht}} + \beta \ln \mathbf{z_{ht}} + \gamma \mathbf{x}_{ht} + \tau_{ht} + \eta_h + \epsilon_{ht}. \tag{5.10}$$

τ_{ht} denote region-time effects – we include a separate set of 90 month dummies for each of 10 broad regions of the Great Britain. η_h denote household fixed effects and $\mathbf{x_{ht}}$ denote time varying household demographics (including age of the youngest child, age of the main shopper, the household's recommended calorie requirement and main shopper employment status).[5]

In our main specification we assume that the coefficients on the basket characteristics are fixed over time. We do this because we estimate equation (5.10) over a period of time where the main changes to the economic landscape are shocks to household income and general food price inflation. In the robustness section we present results where we allow time varying coefficients on the characteristics. This does not change our results qualitatively.

To consistently estimate the parameters in equation (5.10) we require that past, current and future realizations of the right-hand side variables are uncorrelated with the error term. Define $\mathbf{e_h} = (\mathbf{e}_{h1}, ..., \mathbf{e}_{hT})$, $\mathbf{z_h} = (\mathbf{z}_{h1}, ..., \mathbf{z}_{hT})$, $\mathbf{x_h} = (\mathbf{x}_{h1}, ..., \mathbf{x}_{hT})$ and $\boldsymbol{\tau_h} = (\tau_{h1}, ..., \tau_{hT})$; a sufficient condition for identification of the parameters of interest is that the household choice variables

[5] A number of variables entering e and z are bounded between 0 and 1, for these we take the log of 1 plus the variable.

$(\mathbf{e_h}, \mathbf{z_h})$ are strictly exogenous, conditional on the other covariates:

$$\mathbb{E}(\epsilon_{ht}|\mathbf{e_h}, \mathbf{z_h}, \mathbf{x_h}, \boldsymbol{\tau_h}, \eta_h) = 0, \quad t = 1, \ldots, T. \tag{5.11}$$

While we believe that region-time and household fixed effects and time varying household characteristics control for the main potential omitted factors of concern, this is a crucial assumption that we now discuss in further detail.

5.3.3 Identification

We are interested in identifying the causal effect of households' choice variables $(\mathbf{e_{ht}}, \mathbf{z_{ht}})$ on the price per calorie they pay for their grocery basket. Our identification strategy exploits *differential within household* variation in households' shopping choices. The inclusion of household fixed effects, region-time effects and time-varying demographics will help mitigate a number of issues of potential concern.

We do not directly place restrictions on how the disaggregate product prices (p_{isd}) are set (and, in particular, whether the market environment is competitive or oligopolistic). However, we do require market prices to be uncorrelated with the household choice variables $(\mathbf{e_{ht}}, \mathbf{z_{ht}})$, conditional on the household fixed effects, region-time effects and demographics. Market prices are likely to vary over time due to general food price inflation and due to changes in aggregate market conditions feeding into firms' price setting decisions (e.g. firms may put more items on sale during a recession). These price changes may vary regionally. In the UK most supermarkets implement a national pricing policy, following the Competition Commission's investigation into supermarket behavior (Competition Commission (2000)). This means that most regional variation comes from regional variation in supermarket coverage and from differences in temporary price reductions. Such changes will be captured by the region-time effects, τ_{ht} and also by the fact that we control for the availability of food offered on sale, outlined in Section 5.4.1. Similarly, the types of supermarkets located in relatively wealthy areas may set higher prices, and households in such areas may be less inclined to spend time grocery shopping. Purely cross-sectional differences will be controlled for by the household fixed effects, and changes over time (including those that differ across regions) will be absorbed by the region-time effects.

A second possible issue arises if the household varying transaction weights, w_{hisd}, which we use to construct price per calorie, varied in ways other than through, but correlated with, the choice variables of interest. In particular, there may be a variable that influences price paid per calorie that is omitted from the model and that is correlated with those that are included, which would mean that the exogeneity condition (5.11) would not hold. The fact that we include region-time effects and household fixed effects means a problem would arise only if an omitted variable varied over time differentially across households. An example of a possible omitted variable is productivity differences in shopping technology across households within region. For instance, some households may be particularly adept at searching for good deals and consequently may pay less than other households for their groceries. Such households may spend less time shopping and may have preferences that lead them to select different basket characteristics than other households. However, it seems likely that much of the difference in shopping technology would be fixed over time and therefore controlled for by household fixed effects.

Nonetheless it is possible that households' shopping technology and preferences over individual food products may change over time in such a way that is not captured by the included basket characteristics and leads to a lower price per calorie. Two possible reasons for this are changes in household demographics (e.g. the birth of a baby) or the employment status of its members. To control for such changes we include a vector of time-varying household characteristics, including the age of the youngest child, the age of the main shopper and the calorie requirement of the household (see Department of Health (1991)). The inclusion of the household's calorie requirement also captures the potential for economies of scale in grocery purchases, i.e. shopping for more people might allow households to reduce the price that they pay per calorie in ways not captured by the characteristics of the basket, $\mathbf{z_{ht}}$. We also include dummy variables indicating whether the main shopper and head of household work full time or part time. We expect that much of the effect of variation in employment status will be captured by our proxies for shopping effort, but inclusion of these variables will control for any that is not.

Of course in the end we cannot rule out that our estimates are influenced by omitted variable bias, but for this to cause us a problem the source would need to be an omitted variable that varies over time-region differentially within households and that is not captured by demographic transitions.

5.4 Measuring shopping behavior

We measure the price that each household pays for its grocery basket in each period, which we express per calorie, P_{ht} (constructed as described in Section 5.3.2). In 2005-2007 the average nominal price was £1.56 per 1000 calories. By 2010-2012 this had increased by 30p to £1.86. This increase was driven both by changes in the market prices that households faced *and* by changes in the decisions that households made over the characteristics of their basket and their shopping effort. In this section we set out how we measure the household choice variables, $(\mathbf{e_{ht}}, \mathbf{z_{ht}})$. We use these in Section 5.5.1 to separate out the part of the change in price paid per calorie that was due to household behavior: we show that household behavior acted to decrease the price per calorie households paid for their groceries, allowing them to purchase a similar number of calories at lower levels of expenditure.

5.4.1 Shopping effort

An important determinant of the price that households pay for their groceries is how much time and effort they allocate to shopping. For example, the shopper will decide how much time to spend comparing prices and searching for good deals on a shopping trip – the more time she spends comparing prices the less she is likely to pay per calorie for a grocery basket with a given set of characteristics. The shopper must also decide how frequently to shop, and how many different stores to visit. More frequent shopping and visiting more stores provides the opportunity to compare prices across days and retailers, potentially allowing the shopper to find better value products.

This is partly facilitated by the fact that identical products are often sold at different prices in different stores. Kaplan and Menzio (2014a) show that in the US there is a high degree of dispersion in the price at which an identical good is sold across stores, within a given geographic market and period of time. Eden (2013) documents price dispersion across goods sold in supermarkets in Chicago and shows that prices are more dispersed for goods in which there is higher uncertainty about aggregate demand. Aguiar and Hurst (2007) argue that older US households exploit this by both shopping more frequently and spending more time shopping, which allows them to pay less for a fixed basket of groceries than it would cost at average prices. Conversely, it is possible that households may find better deals by making less frequent trips and instead buying a larger share of their basket on each trip. Kaplan and Menzio (2014b) use US time use data to show that employed people spend between 13% and 20% less than unemployed people and scanner data to show that the prices paid by employed workers are 2% higher than those paid by unemployed workers.

We do not directly observe the amount of time households allocate to grocery shopping. We proxy shopping effort using outcome measures from our data. Table 5.4 describes these measures, showing the average value across households in 2005-2007 and 2010-2012, as well as the average within household change and percentage change between these two periods.

The first row of Table 5.4 shows the average number of shopping trips households make per month and the second row shows the average number of separate retailers that they visit. Between 2005-2007 and 2010-2012 households did not change the number of shopping trips that they undertook but they did increase the number of different retailers that they visited. A particularly relevant type of retailer is the discounters; in the third row we report the average share

Table 5.4: *Proxies for shopping effort*

	2005 -2007	2010 -2012	Change	% change
Number of shopping trips (Ntrips)	14.87	14.87	-0.00	-0.00
Number of chains visited (Nstores)	3.70	3.83	0.13	3.44
Share of calories from discounter (DISCOUNTER)	10.24	11.85	1.61	15.67
Share of calories bought on sale (SALE)	24.84	33.93	9.09	36.60
Share of available calories on sale (SALE_AV)	17.19	22.71	5.51	32.06

Notes: The numbers are the mean of each variable in 2005-2007 and 2010-2012 and the average within household change and percentage change. Variable names are shown in brackets. SALE_AV is not a measure of shopping effort; rather we control for it when estimating the price function and, conditional on it, interpret SALE as a measure of shopping effort.

of calories bought from discounters, which increased from 2005-2007 to 2010-2012. Discounters are chains that advertise lower prices compared with other retailers; they are generally less conveniently located and offer a less attractive shopping experience. It is unusual for a household to buy its entire grocery basket at a discounter, because they typically offer a restricted range of products. The share of calories a household purchases at discounter outlets averages 10%. This compares to an average of around 25% in the largest single retailer, Tesco, and over two-thirds in the biggest four supermarkets (Tesco, Asda, Morrisons and Sainsbury's) combined. In the UK the main discounters are Aldi, Iceland, Kwik Save, Lidl and Netto. Prices paid at discounters are typically lower than those paid at other supermarket chains, although much of this is due to differences in the grocery basket composition, meaning that it is important to control for basket characteristics.

Our fourth proxy for shopping effort is designed to capture the amount of time households spend shopping *while in the store*. We measure how intensively households make use of sales as the share of calories they purchase on sale. The idea is that buying a larger than average share of groceries on sale, conditional on basket characteristics, indicates more effort in the shop seeking out the products that the household wants that are on sale. For this interpretation to be valid it is important to account for changes in the number of calories that are available on sale. We therefore control for the share of available calories on sale in the supermarkets that the household visited. Since we also include household fixed effects, this means that the coefficient on the share of calories purchased on sale in the price regression reflects the impact of buying more calories on sale *than the household normally does and holding fixed the share of available calories on sale*. Table 5.4 shows that the share of calories purchased on sale increased substantially from 25% in 2005-2007 to just under 34% in 2010-2012. The share of calories available on sale also increased, but by less - from 17% in 2005-2007 to 23% in 2010-2012. The increase in share of calories available on sale is evident (and of a similar magnitude) across all main food groups.

Note that an important feature of the US grocery market is the availability of coupons that can be collected from newspapers and magazines and can be used to lower the transaction price of specific grocery products. Nevo and Wong (2014) show that, in the US, over the recession increased coupon usage was an important channel through which consumers increased their shopping effort. In contrast, in the UK coupons are not an important feature of the grocery market. Most UK supermarkets do have store loyalty cards. Typically these allow consumers to accumulate points in proportion to their total in store spend, which can be used to lower future grocery bills. For example, the Nectar store card gives customers a point worth 0.5p for every £1 spent in Sainsbury's. These points are collected passively and therefore do not represent increased shopping effort in the way increased coupon usage does in the US market.

5.4.2 Basket characteristics

As well as choosing shopping effort and total calories, households choose the characteristics of their shopping basket, $\mathbf{z_{ht}}$. Basket characteristics include the nutritional characteristics (share of calories from the macronutrients and major food groups, and the amount of the micronutrients) and other characteristics including the share of calories that are bought as budget store brands (i.e. generics) rather than branded products, and package size (to reflect non-linear pricing and bulk discounts). Households may have reduced the price they pay for their groceries without changing the nutritional composition of their calories by adjusting these other characteristics.

Table 5.5 details the nutrient characteristics that we include in $\mathbf{z_{ht}}$. These include the share of non-alcohol calories from each of the macronutrients – protein, saturated fat, unsaturated fat, sugar and non-sugar carbohydrates. All calories are derived from macronutrients (and alcohol), meaning that the shares sum to one. The table shows that between 2005-2007 and 2010-2012, on average, households switched towards carbohydrates (sugar and non-sugar) and unsaturated fat and away from calories from protein and saturated fat. We also include the amount of fibre and salt per 100g in the shopping basket in $\mathbf{z_{ht}}$. Households, on average, have increased the fibre intensity and reduced the salt intensity of their groceries. It is likely that the marginal impact on price paid per calorie of changing nutrients will vary across nutrients because the cost of producing foods with different nutrients varies and because firms might price nutrients differently (for example, Stanley and Tschirhart (1991) find different hedonic prices for nutrients in breakfast cereals). We also control for the nutritional composition of shopping baskets by including in $\mathbf{z_{ht}}$ the share of calories from each of 11 (exhaustive) food groups. Between 2005-2007 and 2010-2012 households, on average, switched towards fruit, grains, poultry and fish, and prepared foods and away from vegetables, red meat and nuts, drinks and alcohol.

We do not have time-use data so do not directly measure how much time households allocated to cooking. However, by controlling for both the nutritional and food group composition of households' grocery baskets, we are able to proxy for the cooking requirement of households' calories (to the extent that cooking times vary across these food groups). For example, if a household switches from purchasing vegetables and raw meats to purchasing processed or prepared foods this indicates a reduction in the required cooking time of its shopping basket. Although we can control for this, we are not able to separately identify how an additional minute of cooking time affects price paid per calorie from the preferences people have over nutrients and food groups.

Table 5.5: *Nutrient characteristics*

Share of calories from:	2005 -2007	2010 -2012	Change	% change
Protein (shr_prot)	14.88	14.76	-0.12	-0.81
Saturated fat (shr_sfat)	14.83	14.59	-0.23	-1.57
Unsaturated fat (shr_ufat)	22.64	22.79	0.15	0.67
Sugar (shr_sug)	22.73	22.82	0.09	0.41
Non-sugar carbohydrates (shr_othcarbs)	24.92	25.03	0.11	0.43
g per 100g of:				
Fibre (fibre)	1.12	1.19	0.07	6.32
Salt (salt)	0.50	0.49	-0.00	-0.10
Share of calories from:				
Fruit (shr_Fruit)	5.08	5.28	0.20	3.86
Vegetables (shr_Veg)	6.97	6.43	-0.54	-7.81
Grains (shr_Grains)	16.40	16.65	0.24	1.48
Dairy (shr_Dairy)	9.53	9.49	-0.04	-0.46
Cheese and fats (shr_CheeseFats)	11.73	11.73	0.01	0.06
Poultry and fish (shr_PoultryFish)	3.09	3.30	0.21	6.87
Red meat and nuts (shr_RedMeatNuts)	8.34	7.84	-0.51	-6.07
Drinks (shr_Drinks)	1.87	1.82	-0.04	-2.36
Prepared sweet (shr_PrepSweet)	19.06	19.53	0.47	2.47
Prepared savory (shr_PrepSavory)	14.78	14.82	0.04	0.30
Alcohol (shr_Alcohol)	3.14	3.11	-0.04	-1.15

Notes: The numbers are mean of each variable in 2005-2007 and 2010-2012 and the average within household change and percentage change. Variable names are shown in brackets.

Table 5.6 details the other (non-nutrient) characteristics we include in $\mathbf{z_{ht}}$. The measure in the first row is the share of calories from budget store brand (or generics). In the UK, there are two types of store brand product: budget and standard. Standard store brands are similar to national brands – they are advertised by the supermarkets, comparably priced and are generally of similar quality to equivalent national brands. In contrast, budget store brands are seldom advertised, are typically sold in plain packaging and are sold for substantially lower prices. The average unit price of budget store brands (across 110 product categories and 16 retailer chains) is just under £2, compared to an average of over £4 for the largest national brand in each product category (Griffith et al. (2014)). Budget store brands are similar to generic brands in the US market. All else equal, it is likely that households value budget store brands less than branded products, and there is evidence that households substitute towards generic products when economic conditions worsen (see Gicheva et al. (2010), Kumcu and Kaufman (2011)). Between 2005-2007 and 2010-2012 households switched to buying a larger share of their calories from generic products.

Griffith et al. (2009) present evidence of strong non-linear pricing in the UK grocery market. Households are able to lower the per calorie price they pay, while keeping other attributes of their shopping basket fixed, by switching to larger pack sizes of the brands they purchase. To capture this we include the share of calories purchased in "big" pack sizes. We define a product as having a "big" pack size if its size is above the median pack size of all transactions involving products belonging to the same brand. The second row of Table 5.6 shows that households switched to buying smaller pack sizes between 2005-2007 and 2010-2012.

Table 5.6: *Other basket characteristics*

Share of calories from:	2005-2007	2010-2012	Change	% Change
Generic products (GEN)	10.92	12.97	2.05	18.75
Big pack sizes (BIG)	32.31	30.86	-1.46	-4.51

Notes: The numbers are mean of each variable in 2005-2007 and 2010-2012 and the average within household change and percentage change. Variable names are shown in brackets.

5.5 Empirical results

In this section we present estimates of the relationship between price paid per calorie and households' choice variables $(\mathbf{e_{ht}}, \mathbf{z_{ht}})$, see equation (5.10). We use the estimates to quantify the contribution that changes in households' behavior made to the change in the average price that they paid for their shopping basket, and we explore the importance of various margins of adjustment. It was by lowering the average price of their shopping baskets that households were able to smooth their calorie purchases over this period; the results in this section show that they did this by increasing their shopping effort, switching to generic products and substituting across nutrients, which, although reducing their utility from consumption, did not adversely impact the nutritional quality of their grocery basket. We also show that the relative importance of these different mechanisms does not differ much across household types.

5.5.1 Estimates of price function

Table 5.7 shows the estimates of the coefficients in equation (5.10). Column (1) shows the estimated coefficients omitting household effects. In column (2) we include household fixed effects. The difference in coefficient estimates is marked. For instance, the absolute value of the sales coefficient more than halves once we include household fixed effects; there are differences in household shopping technology, which leads them to pay a lower price per calorie and that are correlated with their use of sales. A similar change is evident for the other choice variables, underlining the importance of exploiting differential within household changes in behavior. In column (3) we also control for time-varying household characteristics (age of youngest child, age of main shopper, household calorie requirement and employment status). This has much less impact on the coefficient estimates. In what follows we use the coefficient estimates from column (3).

Table 5.7: *Coefficient estimates*

	(1) $\ln(P_{ht})$		(2) $\ln(P_{ht})$		(3) $\ln(P_{ht})$	
ln(Ntrips)	−0.031***	(0.001)	0.021***	(0.001)	0.022***	(0.001)
ln(Nstores)	0.045***	(0.001)	0.010***	(0.001)	0.010***	(0.001)
ln(DISCOUNTER+1)	−0.068***	(0.003)	−0.065***	(0.002)	−0.066***	(0.002)
ln(SALE+1)	−0.348***	(0.003)	−0.143***	(0.003)	−0.141***	(0.003)
ln(SALE_AV+1)	−2.148***	(0.012)	−0.578***	(0.011)	−0.577***	(0.011)
ln(GEN+1)	−1.119***	(0.003)	−0.501***	(0.003)	−0.499***	(0.003)
ln(BIG+1)	−0.467***	(0.003)	−0.218***	(0.003)	−0.216***	(0.003)
ln(shr_sug+1)	0.361***	(0.012)	0.141***	(0.009)	0.142***	(0.009)
ln(shr_sfat+1)	1.941***	(0.014)	1.098***	(0.012)	1.094***	(0.012)
ln(shr_ufat+1)	1.025***	(0.014)	0.379***	(0.011)	0.374***	(0.011)
ln(shr_prot+1)	5.512***	(0.019)	4.073***	(0.015)	4.063***	(0.015)
ln(fibre)	−0.004***	(0.001)	−0.063***	(0.001)	−0.064***	(0.001)
ln(salt)	−0.026***	(0.001)	−0.010***	(0.000)	−0.010***	(0.000)
ln(shr_Fruit+1)	2.402***	(0.010)	1.602***	(0.009)	1.595***	(0.009)
ln(shr_Veg+1)	0.578***	(0.007)	0.459***	(0.006)	0.459***	(0.006)
ln(shr_Dairy+1)	−0.327***	(0.009)	−0.005	(0.008)	−0.005	(0.008)
ln(shr_CheeseFats+1)	−0.554***	(0.010)	−0.249***	(0.008)	−0.245***	(0.008)
ln(shr_RedMeatNuts+1)	−0.549***	(0.010)	−0.084***	(0.008)	−0.080***	(0.008)
ln(shr_PoultryFish+1)	−0.843***	(0.014)	−0.566***	(0.011)	−0.559***	(0.011)
ln(shr_Drinks+1)	1.147***	(0.013)	0.949***	(0.011)	0.948***	(0.011)
ln(shr_PrepSweet+1)	0.333***	(0.007)	0.289***	(0.006)	0.289***	(0.006)
ln(shr_PrepSavory+1)	0.608***	(0.007)	0.657***	(0.006)	0.658***	(0.006)
ln(shr_Alcohol+1)	2.485***	(0.008)	2.163***	(0.008)	2.162***	(0.008)
Region-time effects	Yes		Yes		Yes	
Household fixed effects	No		Yes		Yes	
Time varying hh characteristics	No		No		Yes	

Notes: Estimated with 466,341 observations on 14,694 households' monthly grocery purchases over 2005-2012. Time varying household characteristics include age of the youngest child, the age of the main shopper, calorie requirement of the household and employment status of household main shopper and household head. Standard errors in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

The unconditional correlation between price paid per calorie and number of shopping trips is negative, but Table 5.7 shows that once we control for other choice variables and household fixed effects, the estimated coefficient on number of shopping trips is positive, although small. Conditional on shopping basket characteristics, household caloric requirements and fixed attributes of households, undertaking an additional shopping trip results in a slight increase in price per calorie. This result differs from Aguiar and Hurst (2007), who find that older households pay lower prices because they shop more frequently than other households. Our setting differs in that we focus on within household changes in behavior, rather on cross sectional comparisons. We also find little impact of visiting an additional retailer on price paid per calorie – the coefficient is positive, but as we show below, economically very small. Our other two measures of shopping effort turn out to be more important. Buying a larger share of calories from discounters, all else equal, lowers price paid per calorie. Purchasing more calories on sale, conditional on controlling for how much food is available on sale, leads to a reduction in price paid per calorie. Both of the "other basket characteristics" have the expected coefficient sign: purchasing a higher share of calories from generic products, or switching towards larger pack sizes acts to lower price paid per calorie, all else equal.

The coefficients on the macronutrients (sugar, saturated fat, unsaturated fat and protein) measure the effect of these characteristics on price per calorie relative to the omitted category, non-sugar carbohydrates. Protein is considerably more expensive than the other macronutrients; non-sugar carbohydrates are the cheapest. More fibrous and more salty food acts to lower price per calorie. The food group coefficients capture the effect on price per calorie relative to grains (the omitted category). The coefficients suggest that, all else equal, increasing the share of calories from alcohol and fruit increases price per calorie by the most, and increasing the shares of cheese and fats and poultry and fish lowers price per calorie by the most. Poultry and fish are a relatively expensive source of calories; the negative coefficient for this group is explained by the fact that we control separately for the share of calories from protein in the regression, and they are a relatively cheap source of protein.

5.5.2 Importance of different adjustment mechanisms

In Section 5.2 we showed that households smoothed the amount of calories they purchased over the Great Recession. They did this by acting to reduce the (real) price per calorie of their shopping baskets, allowing them to purchase the same number of calories for less. In this section we use the estimates from the price function to quantify how important each of the choice variables were in allowing them to do this.

Table 5.8 summarizes these results. The average price per calorie households paid increased by 17.7 log points (around 19.4%) between 2005-2007 and 2010-2012. This increase was driven largely by factors outside households' control, such as general food price inflation. Had households not changed their shopping behavior the average price per calorie would have increased by 20.3 log points (around 22.5%). Changes in within household behavior led to a 2.6 log point (approximately a 3.1%) reduction in price paid per calorie. The bottom three rows of Table 5.8 show the contribution made by changes in shopping effort, nutrient characteristics (including food groups) and other characteristics. Increased shopping effort acted to lower the average price paid per calorie by 1.06 log points; changes in the nutrient characteristics acted to lower it by 0.93 log points; changes in the other characteristics of the shopping basket acted to lower price paid by 0.60 log points. All three mechanisms were important in allowing households to smooth their consumption over this period.

Table 5.8: *Changes in log price paid per calorie; estimates from model*

	All households
Change in $\widehat{\ln(P_{ht})}$	17.74
Change in $\widehat{\ln(P_{ht})}$, no behavior	20.34
Change in $\widehat{\ln(P_{ht})}$, due to behavior	-2.59
of which	
shopping effort	-1.06
nutrient characteristics	-0.93
other characteristics	-0.60

Notes: Numbers are the average within household change. Row 1 is change in predicted $\ln(P_{ht})$*. Row 2 is change in predicted* $\ln(P_{ht})$ *holding fixed the choice variables* $(\mathbf{e_{ht}}, \mathbf{z_{ht}})$*. Row 3 is change in predicted* $\ln(P_{ht})$ *holding fixed all variables other than the choice variables* $(\mathbf{e_{ht}}, \mathbf{z_{ht}})$*. All numbers are multiplied by 100.*

Table 5.9: *Contribution of choice variables to change in price paid per calorie*

	Contribution
Shopping effort:	
Number of shopping trips	-0.02
Number of chains visited	0.03
Savings from discounter	-0.09
Savings from sales	-0.97
Total	-1.06
Nutrient characteristics:	
Protein	-0.43
Saturated fat	-0.22
Unsaturated fat	0.05
Sugar	0.01
Fibre	-0.39
Salt	0.06
Fruit	0.28
Vegetables	-0.23
Dairy	0.00
Cheese and fats	-0.00
Poultry and fish	-0.11
Red meat and nuts	0.04
Drinks	-0.04
Prepared sweet	0.11
Prepared savory	0.02
Alcohol	-0.08
Total	-0.93
Other characteristics:	
Share from generic products	-0.84
Share of groceries from big pack sizes	0.24
Total	-0.60
Total	-2.59

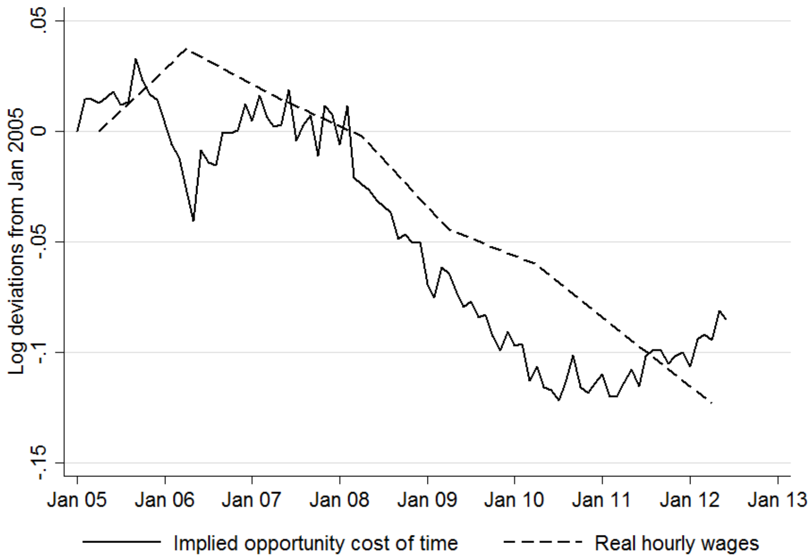
Notes: The table reports the contribution each variable made to the fall in price paid per calorie. The contribution is given by the product of the coefficient in column 3 of Table 5.7 and average change in log of the transformed variable, controlling for fixed effects (multiplied by 100).

In Table 5.9 we present further details of the contribution of changes in each choice variable to the overall 2.6 log point decline in price paid per calorie. The use of sales (holding fixed the amount of calories available on sale) is the most important mechanism that households used. A switch towards buying more calories from less preferred generic products was important in reducing price paid per calorie, leading to a 0.84 log point reduction. Substitution to smaller pack sizes acted to increase price paid per calorie by 0.24 log points.

The reduction in price per calorie through changing the nutritional characteristics was principally due to a switch away from protein, saturated fat and alcohol (all relatively costly per calorie) and towards fibre, non-sugar carbohydrates and vegetables (which are relatively cheap per calorie). Although households changed the nutritional composition of their shopping basket, as we showed in Section 5.2.2, this did not lead to a fall in the overall nutritional quality of the basket for two reasons. First, the reasonably large differences in the relative prices of nutrients means that even small changes in the nutritional balance of the basket can have a considerable impact on its price. Second, households substituted across nutrients and food groups in such a way that the "good" changes offset the "bad", allowing them to maintain the same average nutritional quality as they had purchased prior to the recession. Households switched towards cheaper, less preferred, characteristics (e.g. generic products) and away from more expensive, more preferred, characteristics (e.g. protein, alcohol) in such a way as to maintain the number of calories they were able to purchase and the average nutritional quality of these calories.

We argue that the use of sales (conditional on the availability of products that are offered on sale) is a proxy for effort or time spent shopping. The model we outline in Section 5.3.1 (condition (5.3) in particular) implies that we can use observed changes in households' shopping effort and their grocery purchases to infer how the opportunity cost of time has varied over time. As Aguiar and Hurst (2007) point out, this measure of the opportunity cost of time has the advantage that it allows us to be agnostic about households' behavior in the labor market. Given the functional form we assume for the price function, we can write the opportunity cost of time as $\omega_{ht} = -\alpha \frac{\tilde{P}_{ht} C_{ht}}{1+e_{ht}}$ where $\tilde{P}_{ht}$ is expressed in "real" terms (meaning that variation over time in $\tilde{P}_{ht}$ captures changes in price paid per calorie resulting from changes in household behavior; general food price inflation is removed). The solid line in Figure 5.3 plots the average path of the implied opportunity cost of time over 2005-2012. Over this time period households reduced their real food expenditure, but increased their shopping effort (as measured by our proxy), and this suggests a fall in the opportunity cost of time. As a comparison the dashed line shows real mean gross hourly wages. Our estimate of the opportunity cost of time tracks the cost of time as measured by mean wages reasonably closely.

Figure 5.3: *Implied opportunity cost of time*



Notes: Solid line shows deviations of logged opportunity cost of time from its value in January 2005, after deseasonalising and controlling for fixed effects, and is smoothed using a 7-point moving average. The dashed line plots real hourly wages: mean gross hourly wages from the Annual Survey of Hours and Earnings deflated using the food and drink component of the CPI.

In Section 5.2 we showed that households of all types acted to smooth their calorie purchases over the recession, despite large declines in real food expenditure. We explore whether the ways in which they did this varied across households of different types, both by household composition and household income. Table 5.10 repeats the analysis above for the different household groups. The first three columns show the average change, the change in the absence of any behavioral change and the change due to households' adjustments in behavior.The remaining columns separate the change due to behavior into the contributions made by households' decisions over: shopping effort, nutrient characteristics and other characteristics. Tables C.2 and C.3 in Appendix Section C.1 break this down and provides details of the contribution made by each of the individual choice variables that we include in the price regression.

Households with pre-school children acted to decrease the per calorie price they paid by over 5 log points – more than other household types. Approximately 30% of this was due to increased shopping effort and, in particular, a greater use of sales. Households with young children also switched to buying more of their groceries in the form of generic products. The remaining reduction in the average price per calorie is due to changes in the nutritional composition of their shopping basket, mainly through a fall in the share of calories bought as protein and saturated fat. As shown in Section 5.2.2, this did not lead to a fall in the average nutritional quality of the baskets purchased by these households. Households with school-age children also

Table 5.10: *Changes in log price paid per calorie, by household composition*

	(1)	(2)	(3)	(4)	(5)	(6)
	Change in $\widehat{\ln(P_{ht})}$			Change due to behavior, *of which:*		
	Total	no behavior	due to behavior	Shopping effort	Characteristics: Nutrient	Other
All	17.74	20.34	-2.59	-1.06	-0.93	-0.60
pre-school children	13.98	19.16	-5.19	-1.66	-2.76	-0.77
school aged children	18.57	19.80	-1.23	-1.36	0.37	-0.24
adults	17.74	20.31	-2.57	-1.00	-0.99	-0.59
pensioners	18.13	20.66	-2.53	-0.87	-0.93	-0.73
working high income	16.14	19.73	-3.58	-1.26	-1.56	-0.76
working mid income	17.45	20.03	-2.57	-1.22	-0.79	-0.57
working low income	18.35	20.39	-2.04	-1.14	-0.23	-0.68
unemployed	18.06	20.34	-2.27	-1.15	-1.02	-0.10

Notes: Column 1 is change in predicted $\ln(P_{ht})$*. Column 2 is change in predicted* $\ln(P_{ht})$ *holding fixed the choice variables* $(\mathbf{e_{ht}}, \mathbf{z_{ht}})$*. Column 3 is change in predicted* $\ln(P_{ht})$ *holding fixed all variables other than the choice variables* $(\mathbf{e_{ht}}, \mathbf{z_{ht}})$*; columns 4-6 show the contribution of changes in shopping effort, nutrient characteristics and other characteristics to the change due to behavior shown in column 3. All numbers are multiplied by 100. Household group definitions shown in the notes to Table 5.1.*

reduced the price that they paid per calorie, but by less than households with younger children: the majority of the fall is due to an increase in the use of sales.

Households without children (both pensioner and non-pensioner) changed their behavior in similar ways to each other. The overall effect of changes in their shopping behavior was to reduce the price they paid per calorie by around 2.5 log points. Like other household types, households with no children lowered their price paid per calorie by making greater use of sales, and like households with pre-school children, they switched towards cheaper nutrients and food groups.

The results are similar when we conduct the analysis across households of different income levels. Working households with higher income reduced the price they paid per calorie by the most – by over 3.5 log points. They saved 1.3 log points through greater use of sales. Working households with middle levels of income also increased their use of sales; the big difference between these and higher income households is that the latter group switched more towards cheaper nutrients. Low income working households and unemployed households increased their shopping efforts by similar amounts, but while low income working households substituted more to generic products, unemployed households opted to switch between nutrients to reduce the price they paid per calorie.

Although the relative importance of the different mechanisms varies somewhat, every household type (apart from those with school-age children, who switched to more expensive nutrients) used all three mechanisms to smooth their calorie purchases over this period. Even for those household types for which substitution across nutrient characteristics was particularly important, the average nutritional quality of the baskets of these household types remained stable over this period: households acted to smooth both the quantity and nutritional quality of food purchased over the recessionary period.

5.5.3 Robustness

Functional form of price equation

To check that our results are not driven by the double-log functional form we assume for the price function, we repeat the analysis using an alternative polynomial specification:

$$P_{ht} = a_1 \mathbf{e_{ht}} + a_2 \mathbf{e_{ht}}' \mathbf{e_{ht}} + b_1 \mathbf{z_{ht}} + b_2 \mathbf{z_{ht}}' \mathbf{z_{ht}} + \gamma \mathbf{x}_{ht} + \tau_{ht} + \eta_h + \epsilon_{ht}, \qquad (5.12)$$

maintaining the same exogeneity assumption (5.11). Rather than repeat all tables from Section 5.5 we note that both the baseline and polynomial specification predict approximately a 3% fall in average price paid per calorie due to variation in household behavior and in the first two columns of Table 5.11, for each specification, we report the percentage contribution that each of changes in shopping effort, nutrient characteristics and other characteristics made to this reduction. This shows that both specifications yield similar results.

We also estimate the double-log model letting the coefficients on the basket characteristic, $\mathbf{z}$, vary across the pre, during and post Great Recession time periods. This allows for the possibility that differential inflation across food products may have changed the implicit relative price of characteristics. In the third column of Table 5.11 we summarize the results from this specification. Allowing for time-varying characteristic coefficients yields an even larger impact of household behavior on price paid per calorie; reinforcing our findings. The relative contribution of each channel of adjustment is broadly similar to our baseline model (results available from the authors on request).

Table 5.11: *Changes in log price paid per calorie, alternative specification*

	Specification		
	Double-log (baseline)	Polynomial	Time varying $\mathbf{z}$ coefficients
% change in P_{ht} due to behavior	-3.1	-3.0	-4.8
share due to			
shopping effort	40.8%	45.6%	49.0%
nutrient characteristics	35.8%	34.1%	28.7%
other characteristics	23.1%	20.3%	22.2%

Notes: Row 1 is the percentage change in P_{ht}, holding fixed all variables other than the choice variables ($\mathbf{e_{ht}}, \mathbf{z_{ht}}$). It shows average within household changes. Rows 3-5 show the fraction of the decline that is attributable to each set of choice variables. Column 1 of this table corresponds to the bottom 4 rows of Table 5.8; here the numbers are percentage changes rather than changes in log points.

Food out

Our data are very detailed for food purchased for home consumption, in particular allowing us to measure price and nutrients very accurately. We do not have the same kind of detailed information on purchases of food that is consumed outside the home (e.g. restaurant food and takeaways). However, from the Living Costs and Food Survey (LCFS) we know that although food out (which includes takeaways and food eaten in restaurants) constitutes approximately 36% of total food expenditure, it accounts for only 12-13% of total calories purchased. Therefore, nutritionally, food at home is by far the most important component of households' total food consumption.

We use the data to look at the changes in real expenditure and calories for food at home, which fell by around 6% and 1% respectively - similar changes to those we see in the Kantar data (shown in Table 5.1). Real expenditure and calories from food out both fell by around 10%. However, the LCFS shows that overall expenditure on food (in and out) fell by 7% between 2005-2007 and 2010-2012 and calories fell by just 2%: the pattern of consumption smoothing is evident across total food purchases, not just for food at home.

Table 5.12: *Changes in food at home and food out*

Real expenditure (£ per adult equivalent per month)	2005-2007	2010-2011	Change	% change
Food at home	121.02	114.00	-7.02	-5.8
Food out	70.45	63.76	-6.69	-9.8
Calories (per adult equivalent per day)				
Food at home	2505	2478	-27	-1.1
Food out	381	342	-39	-10.3

Notes: Data from the Living Costs and Food Survey 2005-2011. Real expenditure on food at home is nominal expenditure on food at home deflated by the CPI component for food and drink at home (in 2008 prices). Real expenditure on food out is nominal expenditure on food out deflated by the CPI component for food eaten out (in 2008 prices). Real expenditure is per adult equivalent per month; calories are per adult equivalent per day.

5.6 Summary and conclusions

Aguiar and Hurst (2005) make a convincing case that observed falls in food expenditure at retirement do not translate into falls in consumption. Rather, households increase time spent shopping and in home production to hold their food consumption broadly constant over retirement, in part due to the fall in their opportunity cost of time. Nevo and Wong (2014) and Coibion et al. (2014) show that US households used similar mechanisms to cope with the Great Recession. In this paper we show that in response to *unexpected* worsening in the economic environment households acted to smooth two aspects of their consumption – total calories and their nutritional quality – by increasing their shopping effort and adjusting other aspects of consumption, namely the characteristics of their shopping basket. This provides a further explanation for how households are able to use alternative mechanisms to partially insure themselves against adverse shocks (Blundell et al. (2014)).

We use detailed longitudinal data on grocery purchases that span the period of the Great Recession. Over this period the economic environment deteriorated substantially. Households were subject to depressed real wages, higher unemployment and asset price reductions. At the same time, food prices rose sharply. While some households may have been shielded by the benefit system from the income and asset price shocks associated with the recession, all households faced increases in the price of food relative to the overall price level. We show that households changed their shopping behavior in ways that lowered the average per calorie price of their shopping basket. Spending more time shopping and switching to less preferred characteristics of the shopping basket (which would have made households worse off), nonetheless allowed them to maintain their calorie purchases while reducing their real food expenditure.

The reduction in average price per calorie has raised concern that people have switched to foods of poorer nutritional quality. We show that much of the decline in per calorie spend was driven by margins of change which do not involve altering the nutritional quality of food baskets: households expended more effort shopping (in particular increasing their use of sales) and switched to lower priced generic products. Nevertheless, for most household types, there was substitution towards cheaper nutrients and food groups. Using a single index measure of diet quality we quantify the nutritional importance of these changes and show that the average nutritional quality of food purchases did not materially fall. Our overall conclusion is that households are better able to weather economic turbulence than is suggested by merely looking at their aggregate food expenditure.

Chapter 6

Conclusion

The papers in this thesis all address a particular question concerning how economic agents make decisions in particular settings. They do this by combining economic theory, rigorous microeconometric methods and very detailed micro data on decisions taken by economic agents. The first half to the thesis focuses on the behavior of consumers and firms and how they interact in particular grocery markets. In Chapter 2 we focus on how firm pricing and advertising decisions are made and what their impact is on consumers in the UK market for potato chips. In Chapter 3 we focus on the importance of modeling income effects in random utility models of demand with an application to the UK butter and margarine market. In each case we use the model of the behavior of economic agents in the market to simulate the *ex ante* impact on market equilibria of a policy change (an advertising ban in Chapter 2, tax reform in Chapter 3).

In Chapter 4 we focus on the behavior of innovative multinational firms, modeling where they choose to legally own their patents. Like the preceding chapters, we use this model to study the implications of policy change, studying recent introductions of preferential tax rates for patent income in several countries. Finally, in Chapter 5, we return to the grocery market to consider the mechanisms consumer use to assist them in smoothing their food consumption over times of economic turbulence.

In each case our findings, to some extent, confound common perceptions; for instance we find that an advertising ban may actually lead to higher demand; modeling income effects, even in very small market share goods, can be important for fully capturing consumer welfare effects; new policies aimed at attracting patent income may result in substantial losses in revenue; and evidence that consumers are well able to smooth important aspects of their food consumption, including its nutritional quality, in tough economic conditions. This stands to underline the importance of using rigorous empirical economics to explore the nature of decision making in realistic market settings. With continually expanding computational power and the increasing availability of very rich data on economic choices, there are sure to be growing opportunities for economists to apply microeconometrics to enhance understanding of how firms and consumers make decisions and how their choices will be affected by government interventions.

Appendices

A

Appendix to "The effects of banning advertising in junk food markets"

A.1 Dynamic oligopoly competition in prices and advertising

In Section 2.3 we argue that because i) prices do not directly affect future demand or the evolution of the state variables and ii) we observe the advertising state variables, optimal price are determined by a static equilibrium pricing condition, and for the purposes of considering an advertising ban, we can restrict our attention to this optimality condition. Here we outline a fully dynamic oligopoly game which implies additional optimality conditions (not required in our case) that will characterize equilibria (dynamic) advertising strategies. We abstract from explicitly considering entry and exit for notational simplicity, but as will become clear, identification of marginal costs of products present in the market is independent of whether we allow for entry and exit of firms or products.

Before describing the details of the dynamic oligopoly game, we start by writing the objective function of a firm as a function of strategic variables, prices and advertising expenditures, and the vectors of state variables. The firm owning product (b, s) chooses the product's price, p_{bst}, and advertising expenditures, e_{bt}, for the brand b in each period t. The intertemporal variable profit of firm j at period 0 is:

$$\sum_{t=0}^{\infty} \beta^t \left[\sum_{(b,s)\in F_j} (p_{bst} - c_{bst})\, s_{bs}\left(\mathbf{a_t}, \mathbf{p_t}, \boldsymbol{\zeta_t}\right) M_t - \sum_{b\in B_j} e_{bt}\right], \tag{A.1}$$

where the time t advertising state vector is $\mathbf{a_t}$. As outlined in Section 2.2, we assume $\mathbf{a_t} = \mathcal{A}\left(\mathbf{a_{t-1}}, \mathbf{e_t}\right) = \delta \mathbf{a_{t-1}} + \mathbf{e_t}$, which means that the dimension of the state space remains finite.

We assume that at each period t all firms observe the total market size, M_t, the vector of all firms' marginal costs $\mathbf{c_t}$, and the aggregate demand shocks $\boldsymbol{\zeta_t}$. We denote the information set $\theta_t = (M_t, \mathbf{c_t}, \boldsymbol{\zeta_t})$. We assume that firms form symmetric expectations about future shocks according to the following assumption:

Assumption: Marginal costs and market size follow independent Markov processes such that for all t, $E_t\left[c_{bst+1}\right] = c_{bst}$, $E_t\left[M_{t+1}\right] = M_t$ and $E_t\left[\boldsymbol{\zeta_{t+1}}\right] = \boldsymbol{\zeta_t}$.

We follow the majority of the empirical literature by restricting our attention to pure Markov strategies (see, inter alia, Ryan (2012), Sweeting (2013) and Dubé et al. (2005)). This restricts firms' strategies to depend only on payoff relevant state variables, $(\mathbf{a_{t-1}}, \theta_t)$. For each firm j, a Markov strategy σ_j is a mapping between the state variables $(\mathbf{a_{t-1}}, \theta_t)$, and the firm j decisions $\left\{p_{bst}\right\}_{(b,s)\in F_j}, \left\{e_{bt}\right\}_{b\in B_j}$, which consist of choosing prices and advertising expenditures for the firm's own products and brands ($\sigma_j\left(\mathbf{a_{t-1}}, \theta_t\right) = (\left\{p_{bst}\right\}_{(b,s)\in F_j}, \left\{e_{bt}\right\}_{b\in B_j})$).

There is no guarantee that a Markov Perfect Equilibrium (MPE) in pure strategies of this dynamic game exists. In a discrete version of this game, existence of a symmetric MPE in pure strategies follows from the arguments in Doraszelski and Satterthwaite (2003, 2010), provided that we impose an upper bound on advertising strategies. Ericson and Pakes (1995) and Doraszelski and Satterthwaite (2003) provide general conditions for the existence of equilibria in similar games, but as our model set up differs, the conditions cannot be directly applied in our case. Therefore we assume the technical conditions for the existence of a subgame perfect Markov Perfect Equilibrium of this game are satisfied, and below we use necessary conditions to characterize an equilibrium (Maskin and Tirole (2001)). However, we do not need to assume that an equilibrium is unique, and indeed it is perfectly possible that this game has multiple equilibria.

In this dynamic oligopoly game, each firm j makes an assumption on the competitors' strategy profiles denoted σ_{-j}, where $\sigma_{-j}\left(\mathbf{a_{t-1}}, \theta_t\right) = (\sigma_1\left(\mathbf{a_{t-1}}, \theta_t\right), .., \sigma_{j-1}\left(\mathbf{a_{t-1}}, \theta_t\right), \sigma_{j+1}\left(\mathbf{a_{t-1}}, \theta_t\right), .., \sigma_J\left(\mathbf{a_{t-1}}, \theta_t\right))$. Equilibrium decisions are generated by a value function, $\pi_j^*(.,.)$, that satisfies the following Bellman equation

$$\pi_j^*\left(\mathbf{a_{t-1}},\theta_t\right)=\max_{\{p_{bst}\}_{(b,s)\in F_j},\{e_{bt}\}_{b\in B_j}}\Bigg\{\sum_{(b,s)\in F_j}(p_{bst}-c_{bst})s_{bs}\left(\mathbf{a_t},\mathbf{p_t},\boldsymbol{\zeta_t}\right)M_t-\sum_{b\in B_j}e_{bt}$$
$$+\beta E_t\left[\pi_j^*\left(\mathbf{a_t},\theta_{t+1}\right)\right]\Bigg\},$$

where

$$\mathbf{a_t}=(a_{1t},..,a_{Bt})=\left(\mathcal{A}\left(a_{1t-1},e_{1t}\right),..,\mathcal{A}\left(a_{Bt-1},e_{Bt}\right)\right).$$

and where $\pi_j^*\left(\mathbf{a_t},\theta_{t+1}\right)$ is the next period discounted profit of firm j, given the vector of future advertising states $\mathbf{a_t}=(a_{1t},..,a_{Bt})$. The Bellman equation is conditional on a specific competitive strategy profile σ_{-j}. A MPE is then a list of strategies, σ_j^* for $j=1,..,J$, such that no firm deviates from the action prescribed by σ_j^* in any subgame that starts at some state $(\mathbf{a_{t-1}},\theta_t)$.

Assuming that the technical conditions for the profit function to be differentiable in price and have a single maximum are satisfied, we can use the firstorder conditions of firm j profit with respect to prices for each $(b,s)\in F_j$:

$$s_{bs}\left(\mathbf{a_t},\mathbf{p_t},\boldsymbol{\zeta_t}\right)+\sum_{(b',s')\in F_j}\left(p_{b's't}-c_{b's't}\right)\frac{\partial s_{b's'}\left(\mathbf{a_t},\mathbf{p_t},\boldsymbol{\zeta_t}\right)}{\partial p_{bst}}=0. \qquad \text{(A.2)}$$

We can identify price-cost margins using the condition (A.2) provided this system of equations is invertible, which will be the case if goods are "connected substitutes" as in Berry and Haile (2014). Another set of conditions for the optimal choice of advertising flows exists and characterizes the equilibrium relationship between advertising flows, prices and all state variables including past advertising. We however do not need to use such a condition for identifying marginal costs since the price first order conditions are sufficient. Thus, we do not need to impose differentiability of the profit function with respect to advertising, nor continuity, we only need to use the necessary firstorder condition on price, which depends on the observed state vector $\mathbf{a_t}$. In addition, if we allowed for entry and exit of firms we still would be able to identify marginal costs using equation (A.2); entry and exit would change optimal advertising expenditures and the set of the firms in the market (both of which we observe), but it would not change the form of the price first order condition for active firms.

As shown by Dubé et al. (2005) and Villas-Boas (1993), this type of dynamic game can give rise to alternating strategies or pulsing strategies in advertising, corresponding to each MPE profile σ. However, the identification of marginal costs, c_{bst}, does not depend on the equilibrium value function π_j^* for a given level of observed optimal prices and advertising $(\mathbf{p_t},\mathbf{e_t})$. Marginal costs will depend on equilibrium strategies only through observed prices and advertising decisions, and will simply be the solution of the system of equations (A.2). Therefore we can identify marginal costs without making assumptions about the uniqueness of dynamic equilibria, whether firms' value function are differentiable, or whether the same equilibria is played in each market.

A.2 Willingness to pay for healthiness

Define the nutrient characteristic x_b such that a higher value corresponds to lower nutritional quality and consider the willingness to pay for a change in x_b that *reduces* product unhealthiness. This is given by

$$WTP_{ibt} = \frac{\partial \bar{v}_{ibst} / \partial x_b}{\partial \bar{v}_{ibst} / \partial p_{bst}} = \frac{\psi_{0d} + \psi_{1d} a_{bt}}{\alpha_{0i} + \alpha_{1d} a_{bt}}. \tag{A.3}$$

If consumers positively value income and healthiness, the marginal effect of price and of the nutrient characteristic in the payoff function will be negative ($\alpha_{0i} + \alpha_{1d} a_{bt} < 0$ and $\psi_{0d} + \psi_{1d} a_{bt} < 0$), and the willingness to pay for a healthier product (a decrease in the nutrient characteristic) will be positive. However, whether the willingness to pay for healthiness will decrease with advertising will depend on the relative signs and magnitudes of the interactions between advertising and price and the nutrient characteristic.

A.3 Expected utility under characteristics view of advertising

Our model specification leads, under the characteristic view of advertising, to expected utility given (up to an additive constant) by:

$$\begin{aligned} W_i(\mathbf{a_t}, \mathbf{p_t}) = \ln \Bigg[\sum_{(b \neq 0, s \neq 0) \in \Omega_\kappa} \exp \Big[& \alpha_{0i} p_{bst} + \psi_{0d} x_b + \eta_d z_{bs} + \xi_{ib} + \big[\lambda_i a_{bt} + \rho_i \Big(\textstyle\sum_{l \neq b} a_{lt} \Big) \\ & + \alpha_{1d} a_{bt} p_{bst} + \psi_{1d} a_{bt} x_b \big] \Big] + \exp\left[\zeta_{d0t}\right] \Bigg] \end{aligned}$$

An alternative to our model specification is:

$$\begin{aligned} \tilde{\bar{v}}_{ibst} &= \alpha_{0i} p_{bst} + \psi_{0d} x_b + \eta_d z_{bs} + \xi_{ib} + \left[\tilde{\lambda}_i a_{bt} + \alpha_{1d} a_{bt} p_{bst} + \psi_{1d} a_{bt} x_b \right] + \epsilon_{ibst} \\ \tilde{\bar{v}}_{i00t} &= \zeta_{d0t} + \tilde{\rho}_i \Big(\textstyle\sum_l a_{lt} \Big) + \epsilon_{i00t}. \end{aligned} \tag{A.4}$$

Note:

$$\widetilde{\widetilde{v}}_{ibst} - \widetilde{\widetilde{v}}_{i00t} = \alpha_{0i}p_{bst} + \psi_{0d}x_b + \eta_d z_{bs} + \xi_{ib} + \left[\widetilde{\lambda}_i a_{bt} - \widetilde{\rho}_i \left(\sum\nolimits_l a_{lt}\right) + \alpha_{1d}a_{bt}p_{bst} + \psi_{1d}a_{bt}x_b\right] - \zeta_{d0t} + (\epsilon_{ibst} - \epsilon_{i00t})$$

Setting $\widetilde{\lambda}_i = \lambda_i - \rho_i$ and $\widetilde{\rho}_i = -\rho_i$ shows that $\widetilde{\widetilde{v}}_{ibst} - \widetilde{\widetilde{v}}_{i00t} = \bar{v}_{ibst} - \bar{v}_{i00t}$ meaning that the alternative specification yields observationally equivalent demand to our main specification.

However, expected utility under equation (A.4) is given by

$$\begin{aligned}
\widetilde{W}_i(\mathbf{a_t}, \mathbf{p_t}) = {} & \ln\Bigg[\sum\nolimits_{(b\neq 0, s\neq 0)\in\Omega_\kappa} \exp\Big[\alpha_{0i}p_{bst} + \psi_{0d}x_b + \eta_d z_{bs} + \xi_{ib} + \left[\widetilde{\lambda}_i a_{bt} + \alpha_{1d}a_{bt}p_{bst}\right.\\
& \left. + \psi_{1d}a_{bt}x_b\right]\Big] + \exp\left[\zeta_{i0t} + \widetilde{\rho}_i\left(\sum\nolimits_l a_{lt}\right)\right]\Bigg]\\
= {} & \ln\Bigg[\sum\nolimits_{(b\neq 0, s\neq 0)\in\Omega_\kappa} \exp\Big[\alpha_{0i}p_{bst} + \psi_{0d}x_b + \eta_d z_{bs} + \xi_{ib} + \left[\widetilde{\lambda}_i a_{bt} + \alpha_{1d}a_{bt}p_{bst}\right.\\
& \left. + \rho_i\left(\sum\nolimits_{l\neq b} a_{lt}\right) + \psi_{1d}a_{bt}x_b\right]\Big] + \exp\left[\zeta_{i0t}\right]\Bigg] - \rho_i\left(\sum\nolimits_l a_{lt}\right)\\
= {} & W_i(\mathbf{a_t}, \mathbf{p_t}) - \rho_i \sum\nolimits_l a_{lt}
\end{aligned}$$

Therefore the two specifications, giving rise to identical demand, lead to different welfare conclusions. Under the characteristic view of advertising welfare is sensitive to whether competitor advertising is included in inside product utilities or whether total advertising is included in outside option utility.

A.4 Additional results

A.4.1 Coefficients estimates

Table A.1: *Coefficient estimates for food at home - part 1*

		No kids, high inc., high sk.	No kids, medium inc., high sk.	No kids, low inc., high sk.	No kids, high-medium inc., low sk.	No kids, low inc., low sk.	Pensioners
Random coefficients							
Means	Price	0.2271	0.5916	0.6273	0.7101	0.5751	0.6511
		0.0611	0.0563	0.0537	0.0587	0.0485	0.0656
	Brand advertising	-0.1080	-0.1289	-0.1172	-0.1331	-0.0831	-0.0909
		0.0161	0.0197	0.0190	0.0217	0.0160	0.0221
	Competitor advertising	0.0014	0.0094	-0.0012	-0.0042	0.0036	-0.0011
		0.0028	0.0037	0.0035	0.0040	0.0031	0.0038
Std. deviations	Price	0.4306	0.4160	0.3410	0.3144	0.3969	0.3003
		0.0271	0.0241	0.0233	0.0228	0.0239	0.0240
	Brand advertising	0.0510	0.0824	0.0743	0.0610	0.0583	0.0721
		0.0030	0.0040	0.0044	0.0044	0.0028	0.0054
	Competitor advertising	0.0231	0.0242	0.0192	0.0181	0.0169	0.0224
		0.0008	0.0009	0.0008	0.0011	0.0006	0.0011
	Walkers	1.1154	0.9032	0.9763	0.8564	1.0295	1.3670
		0.0511	0.0454	0.0412	0.0499	0.0437	0.0732
Fixed coefficients							
	Size	0.0186	0.0245	0.0212	0.0252	0.0227	0.0212
		0.0008	0.0010	0.0010	0.0012	0.0009	0.0011
	Size squared	-0.0202	-0.0241	-0.0200	-0.0255	-0.0217	-0.0208
		0.0011	0.0013	0.0012	0.0015	0.0011	0.0015
	Price*Brand advertising	0.0299	0.0267	0.0344	0.0243	0.0247	0.0250
		0.0027	0.0033	0.0032	0.0038	0.0027	0.0036
	Health cost*Brand advertising	0.0032	0.0055	0.0033	0.0051	0.0020	0.0022
		0.0010	0.0012	0.0011	0.0013	0.0010	0.0014
	Pringles	-1.0651	-1.3695	-1.1231	-1.2253	-1.1006	-1.1834
		0.0965	0.1241	0.1200	0.1378	0.1049	0.1493
	Walkers Regular	-0.0760	0.1166	0.4646	0.4062	0.2585	-0.2125
		0.0920	0.1085	0.1041	0.1211	0.0862	0.1324
	Walkers Sensations	-1.6408	-2.0164	-2.0248	-1.9241	-2.3286	-3.0195
		0.0777	0.1060	0.1132	0.1202	0.1038	0.1662
	Walkers Doritos	-1.7287	-1.7723	-1.9677	-1.9248	-2.2511	-3.2387
		0.0757	0.0873	0.1000	0.1083	0.0865	0.1551
	Walkers Other	-0.1811	-0.0497	0.2828	-0.2166	0.0763	-0.3887
		0.0707	0.0860	0.0877	0.1029	0.0732	0.1108
	KP	-0.7540	-0.9515	-0.5889	-0.5844	-0.7480	-1.2917
		0.0598	0.0754	0.0727	0.0843	0.0627	0.0909
	Golden Wonder	-2.8921	-2.9833	-2.3722	-2.1121	-2.4445	-2.4069
		0.1095	0.1427	0.1118	0.1247	0.0985	0.1204
	Asda	-2.6923	-2.8193	-2.7785	-2.8972	-2.5329	-3.5361
		0.0817	0.0996	0.1009	0.1264	0.0802	0.1399
	Tesco	-2.3669	-2.2055	-2.3265	-2.3282	-2.2447	-2.2236
		0.0737	0.0853	0.0892	0.1057	0.0750	0.0962
	Outside Option	4.3094	4.4588	3.7737	4.2139	4.3027	4.2034
		0.2132	0.2826	0.2694	0.3174	0.2417	0.2941
	2010	0.0162	0.0616	0.0411	0.0675	0.0982	0.0480
		0.0433	0.0614	0.0559	0.0631	0.0494	0.0596
	q2	-0.1507	-0.1607	-0.1930	-0.1408	-0.1501	-0.2033
		0.0574	0.0758	0.0726	0.0851	0.0647	0.0739
	q3	-0.0891	-0.0693	-0.1195	-0.0106	-0.1027	-0.1615
		0.0521	0.0689	0.0663	0.0775	0.0589	0.0686
	q4	-0.1916	-0.1676	-0.2535	-0.1334	-0.1743	-0.1643
		0.0623	0.0819	0.0791	0.0920	0.0706	0.0837

Notes: Each column represents a separate estimation (one for each consumer type group), standard errors are reported below coefficient estimates. Random coefficients have normal distributions except for the price coefficient which is log normal.

Table A.2: *Coefficient estimates for food at home - part 2*

		Kids, high inc., high sk.	Kids, medium inc., high sk.	Kids, low inc., high sk.	Kids, high-med inc., low sk.	Kids, low inc., low sk.
Random coefficients						
Means	Price	0.4413	0.6085	0.3508	0.4810	0.5492
		0.0482	0.0477	0.0869	0.0515	0.0516
	Brand advertising	-0.1178	-0.1330	-0.1640	-0.0774	-0.1359
		0.0143	0.0164	0.0233	0.0147	0.0155
	Competitor advertising	0.0011	0.0028	0.0010	0.0055	0.0058
		0.0027	0.0032	0.0045	0.0031	0.0031
Std. deviations	Price	0.2298	0.2845	0.2795	0.2935	0.2710
		0.0141	0.0156	0.0256	0.0190	0.0156
	Brand advertising	0.0412	0.0527	0.0775	0.0572	0.0385
		0.0028	0.0039	0.0061	0.0034	0.0025
	Competitor advertising	0.0156	0.0171	0.0129	0.0204	0.0117
		0.0007	0.0006	0.0015	0.0008	0.0009
	Walkers	0.9638	0.8614	0.9771	1.0559	0.7748
		0.0416	0.0700	0.0700	0.0510	0.0432
Fixed coefficients						
	Size	0.0222	0.0225	0.0229	0.0234	0.0218
		0.0008	0.0009	0.0013	0.0009	0.0009
	Size squared	-0.0211	-0.0198	-0.0218	-0.0217	-0.0176
		0.0010	0.0011	0.0016	0.0011	0.0011
	Price*Brand advertising	0.0272	0.0357	0.0299	0.0257	0.0296
		0.0024	0.0028	0.0041	0.0026	0.0028
	Health cost*Brand advertising	0.0053	0.0044	0.0060	0.0018	0.0052
		0.0008	0.0010	0.0013	0.0009	0.0009
	Pringles	-1.3287	-1.1682	-1.3315	-0.6425	-0.9136
		0.0893	0.1044	0.1514	0.0896	0.0986
	Walkers Regular	0.1714	0.5298	0.9426	0.8034	0.7943
		0.0810	0.0891	0.1156	0.0865	0.0798
	Walkers Sensations	-1.6247	-1.7131	-1.7191	-1.5261	-2.3419
		0.0774	0.0949	0.1295	0.0919	0.1218
	Walkers Doritos	-1.5155	-1.4264	-1.4439	-1.3567	-1.5015
		0.0665	0.0755	0.1064	0.0793	0.0746
	Walkers Other	0.1090	0.5445	0.3951	0.7308	0.5595
		0.0678	0.0781	0.1052	0.0758	0.0722
	KP	-0.5011	-0.3116	-0.2540	-0.0554	-0.1713
		0.0537	0.0615	0.0870	0.0568	0.0598
	Golden Wonder	-2.8624	-3.1317	-2.6873	-2.3288	-2.1211
		0.1109	0.1440	0.1643	0.1013	0.0965
	Asda	-2.5479	-2.2133	-1.9949	-2.1847	-2.0091
		0.0760	0.0783	0.1065	0.0773	0.0765
	Tesco	-1.8913	-1.9214	-1.8338	-1.7905	-1.8217
		0.0648	0.0747	0.1062	0.0716	0.0753
	Outside Option	4.4781	4.0970	4.2007	4.9210	4.6381
		0.2060	0.2478	0.3524	0.2392	0.2430
	2010	-0.0516	-0.0142	0.0860	0.0414	-0.0689
		0.0410	0.0528	0.0755	0.0483	0.0492
	q2	-0.2114	-0.1716	-0.1685	-0.0661	-0.0887
		0.0559	0.0663	0.0942	0.0659	0.0656
	q3	-0.1585	-0.0745	0.0209	-0.0868	-0.0764
		0.0506	0.0600	0.0852	0.0595	0.0596
	q4	-0.2118	-0.0679	-0.0127	-0.0737	-0.2133
		0.0603	0.0727	0.1016	0.0706	0.0712

Notes: Each column represents a separate estimation (one for each consumer type group), standard errors are reported below coefficient estimates. Random coefficients have normal distributions except for the price coefficient which is log normal.

Table A.3: *Coefficient estimates for food on-the-go - part 1*

		No kids, high inc., high sk.	No kids, medium inc., high sk.	No kids, low inc., high sk.	No kids, high-medium inc., low sk.	No kids, low inc., low sk.	Pensioners
Random coefficients							
Means	Price	1.7746	2.4805	1.9629	2.2737	2.4293	1.7461
		0.1090	0.0752	0.1307	0.0915	0.0870	0.2024
	Brand advertising	0.0062	-0.1089	-0.1008	-0.1103	-0.0786	0.0162
		0.0241	0.0326	0.0334	0.0332	0.0354	0.0440
	Competitor advertising	-0.0117	0.0066	-0.0056	-0.0109	-0.0093	-0.0032
		0.0041	0.0054	0.0057	0.0064	0.0058	0.0070
Std. deviations	Price	0.3723	0.3202	0.2350	0.4213	0.2539	0.3053
		0.0376	0.0240	0.0310	0.0318	0.0246	0.0532
	Brand advertising	0.0962	0.1161	0.1015	0.0904	0.1233	0.0804
		0.0050	0.0062	0.0060	0.0055	0.0084	0.0079
	Competitor advertising	0.0341	0.0325	0.0429	0.0431	0.0311	0.0410
		0.0018	0.0016	0.0023	0.0021	0.0023	0.0023
Fixed coefficients							
	Size	0.1939	0.2572	0.1815	0.1140	0.1867	0.1044
		0.0125	0.0175	0.0185	0.0192	0.0205	0.0260
	Size squared	-2.0349	-2.5178	-1.8449	-0.9922	-1.6177	-1.2951
		0.1209	0.1726	0.1778	0.1824	0.1915	0.2509
	Price*Brand advertising	-0.3086	-0.0599	-0.1394	-0.0864	-0.1286	-0.0983
		0.0327	0.0419	0.0448	0.0436	0.0451	0.0516
	Health cost*Brand advertising	0.0086	0.0109	0.0123	0.0103	0.0065	0.0011
		0.0013	0.0020	0.0020	0.0021	0.0020	0.0030
	Walkers Regular	0.2376	0.6036	0.6746	0.7684	0.7010	0.3476
		0.0917	0.1283	0.1335	0.1332	0.1385	0.1620
	Walkers Sensations	-1.5061	-1.6311	-2.0135	-2.9646	-2.5863	-2.8180
		0.1038	0.1748	0.2009	0.3626	0.2901	0.3473
	Walkers Doritos	-1.9344	-1.5706	-1.4550	-1.6256	-1.0126	-1.9177
		0.1054	0.1307	0.1319	0.1573	0.1321	0.2098
	Walkers Other	-0.1725	-0.0132	-0.2894	-0.3363	0.0337	-0.6659
		0.0720	0.1072	0.1108	0.1209	0.1211	0.1626
	KP	-0.1119	-0.8062	-0.4331	-0.3035	0.0257	-0.5781
		0.0844	0.1318	0.1311	0.1363	0.1315	0.1832
	Golden Wonder	-2.5333	-2.2700	-2.5237	-2.3327	-2.1063	-2.2612
		0.0998	0.1343	0.1495	0.1437	0.1381	0.1646
	Outside Option	3.8908	3.7390	3.5501	-0.1003	2.2044	2.0971
		0.3822	0.5445	0.5641	0.5990	0.5778	0.7208
	2010	-0.2027	-0.4600	-0.0587	0.0503	0.0384	0.0870
		0.0735	0.1063	0.1104	0.1189	0.1080	0.1285
	q2	-0.3135	-0.4052	-0.1186	-0.2188	-0.2797	0.1305
		0.0796	0.1113	0.1144	0.1259	0.1117	0.1321
	q3	-0.2111	-0.7334	-0.2190	-0.2115	-0.3381	0.0443
		0.0782	0.1114	0.1125	0.1254	0.1121	0.1330
	q4	-0.2457	-0.4391	-0.1675	-0.0833	-0.1128	0.0285
		0.0912	0.1281	0.1313	0.1464	0.1310	0.1550

Notes: Each column represents a separate estimation (one for each consumer type group), standard errors are reported below coefficient estimates. Random coefficients have normal distributions except for the price coefficient which is log normal.

Table A.4: *Coefficient estimates for food on-the-go - part 2*

		Kids, high inc., high sk.	Kids, medium inc., high sk.	Kids, low inc., high sk.	Kids, high-med inc., low sk.	Kids, low inc., low educ	Kid purchaser
Random coefficients							
Means	Price	1.8951	2.1390	1.3186	2.1306	1.4332	1.5931
		0.1234	0.0971	0.3044	0.0963	0.2078	0.2070
	Brand advertising	0.0177	-0.0122	0.0594	-0.1194	0.0493	-0.1116
		0.0253	0.0299	0.0441	0.0295	0.0338	0.0524
	Competitor advertising	-0.0061	-0.0022	-0.0108	0.0015	-0.0090	-0.0079
		0.0044	0.0054	0.0073	0.0053	0.0057	0.0090
Std. deviations	Price	0.2490	0.4416	0.3528	0.3066	0.2227	0.4316
		0.0354	0.0462	0.0810	0.0263	0.0465	0.0815
	Brand advertising	0.0794	0.0730	0.1150	0.1431	0.0904	0.1057
		0.0044	0.0052	0.0100	0.0073	0.0057	0.0079
	Competitor advertising	0.0315	0.0318	0.0371	0.0422	0.0405	0.0435
		0.0023	0.0022	0.0027	0.0024	0.0022	0.0038
Fixed coefficients							
	Size	0.2274	0.1615	0.2473	0.2103	0.0702	0.0919
		0.0160	0.0171	0.0238	0.0162	0.0208	0.0302
	Size squared	-2.3861	-1.5996	-2.5982	-2.0731	-0.8988	-0.9626
		0.1587	0.1656	0.2262	0.1574	0.1985	0.2976
	Price*Brand advertising	-0.2414	-0.1271	-0.4051	-0.0385	-0.2236	-0.2102
		0.0383	0.0405	0.0655	0.0357	0.0449	0.0638
	Health cost*Brand advertising	0.0073	0.0065	0.0088	0.0083	0.0040	0.0134
		0.0015	0.0018	0.0027	0.0017	0.0022	0.0034
	Walkers Regular	0.5664	0.2656	-0.0597	0.7684	0.1511	0.9275
		0.1062	0.1252	0.1795	0.1155	0.1480	0.2267
	Walkers Sensations	-1.4374	-1.5957	-1.8681	-1.9429	-1.7566	-1.8593
		0.1385	0.1744	0.1891	0.1768	0.1985	0.3416
	Walkers Doritos	-0.7971	-1.2166	-1.0120	-1.0667	-1.0787	-1.1954
		0.0942	0.1223	0.1318	0.1090	0.1377	0.2289
	Walkers Other	-0.0288	-0.1656	-0.1453	0.0642	-0.4004	-0.3712
		0.0887	0.1035	0.1294	0.0971	0.1250	0.1947
	KP	0.2608	0.0878	-0.6124	-0.0775	0.1174	-0.1754
		0.1001	0.1166	0.1644	0.1114	0.1387	0.2225
	Golden Wonder	-2.3707	-2.2658	-2.8580	-2.6030	-1.4742	-1.6869
		0.1313	0.1376	0.1874	0.1381	0.1183	0.1982
	Outside Option	4.8573	2.3523	5.4420	3.1708	2.4279	2.2514
		0.4684	0.5203	0.7023	0.4937	0.5579	0.8698
	2010	-0.0813	0.0518	0.2424	0.1116	0.0787	0.2156
		0.0792	0.1003	0.1436	0.0940	0.1061	0.1705
	q2	-0.0606	-0.1374	-0.0378	-0.0319	-0.0006	-0.0770
		0.0856	0.1040	0.1427	0.1027	0.1160	0.1903
	q3	-0.0124	-0.1329	-0.2747	0.0199	-0.1726	-0.2395
		0.0853	0.1033	0.1379	0.1017	0.1117	0.1861
	q4	0.0713	-0.0094	-0.0336	-0.0439	-0.1232	-0.3923
		0.0988	0.1203	0.1650	0.1170	0.1307	0.2132

Notes: Each column represents a separate estimation (one for each consumer type group), standard errors are reported below coefficient estimates. Random coefficients have normal distributions except for the price coefficient which is log normal.

A.4.2 Mean market price elasticities

Table A.5: *Own and cross price elasticities*

Selected food at home products

	Walkers Regular 150-300g	Walkers Regular 300g+	Walkers Sensations 150-300g	Walkers Sensations 300g+	Walkers Doritos 150-300g	Walkers Doritos 300g+	Walkers Other 150-300g	Walkers Other 300g+	Walkers Other <150g	Pringles 150-300g	Pringles 300g+
Walkers Regular:150-300g	-1.4494	0.4126	0.0149	0.0340	0.0283	0.0559	0.0470	0.1116	0.1900	0.0196	0.0701
Walkers Regular:300g+	0.0561	-2.1529	0.0105	0.0347	0.0204	0.0604	0.0318	0.0901	0.2219	0.0116	0.0719
Walkers Sensations:150-300g	0.0571	0.2687	-1.9769	0.0503	0.0333	0.0658	0.0442	0.1070	0.1888	0.0121	0.0479
Walkers Sensations:300g+	0.0412	0.2792	0.0167	-3.2778	0.0257	0.0692	0.0324	0.0906	0.2118	0.0088	0.0496
Walkers Doritos:150-300g	0.0625	0.3072	0.0188	0.0447	-1.7350	0.0691	0.0446	0.1092	0.1956	0.0141	0.0570
Walkers Doritos:300g+	0.0447	0.3260	0.0141	0.0456	0.0259	-2.9895	0.0330	0.0932	0.2230	0.0101	0.0597
Walkers Other:<150g	0.0705	0.3300	0.0173	0.0394	0.0305	0.0609	-1.6455	0.1212	0.2080	0.0156	0.0623
Walkers Other:150-300g	0.0603	0.3352	0.0154	0.0406	0.0273	0.0630	0.0438	-2.0891	0.2212	0.0135	0.0636
Walkers Other:300g+	0.0412	0.3293	0.0117	0.0408	0.0206	0.0636	0.0311	0.0912	-2.9505	0.0094	0.0624
Pringles:150-300g	0.0530	0.2283	0.0087	0.0201	0.0176	0.0352	0.0288	0.0693	0.1213	-1.3501	0.1066
Pringles:300g+	0.0326	0.2394	0.0066	0.0216	0.0133	0.0383	0.0205	0.0583	0.1441	0.0181	-2.3205
Outside option	0.0265	0.1138	0.0067	0.0141	0.0119	0.0220	0.0178	0.0408	0.0660	0.0110	0.0381

Food-on-the-go products

	Walkers Regular 34.5g	Walkers Regular 50g	Walkers Sensations	Walkers Doritos	Walkers Other <30g	Walkers Other 30g+	KP	GW <40g	GW 40g+	Other <40g	Other 40g+
Walkers Regular:34.5g	-3.4281	0.3621	0.0259	0.0907	0.1908	0.2277	0.4456	0.0386	0.0092	0.2474	0.0730
Walkers Regular:50g	1.1199	-5.5032	0.0273	0.0920	0.1789	0.2441	0.4412	0.0346	0.0101	0.2435	0.0816
Walkers Sensations:40g	0.4802	0.1743	-4.0349	0.1182	0.1625	0.2275	0.5397	0.0747	0.0254	0.3481	0.1046
Walkers Doritos:40g	0.6438	0.2212	0.0439	-4.2025	0.1818	0.2438	0.5236	0.0667	0.0189	0.3376	0.1042
Walkers Other:<30g	0.7445	0.2343	0.0360	0.1031	-3.8489	0.2692	0.5196	0.0588	0.0147	0.3151	0.0968
Walkers Other:30g+	0.7228	0.2585	0.0393	0.1099	0.2207	-4.8892	0.5322	0.0499	0.0152	0.3070	0.0985
KP:50g	0.6390	0.2156	0.0454	0.1111	0.1905	0.2431	-3.6772	0.0643	0.0179	0.3264	0.1033
Golden Wonder:<40g	0.4348	0.1401	0.0510	0.1148	0.1654	0.1822	0.5057	-2.6589	0.0330	0.4043	0.1207
Golden Wonder:40g+	0.3619	0.1397	0.0535	0.1106	0.1452	0.1872	0.4650	0.0985	-3.7928	0.3540	0.1196
Other:<40g	0.5753	0.1945	0.0455	0.1157	0.1867	0.2313	0.5272	0.0809	0.0218	-3.5217	0.1174
Other:40g+	0.5602	0.2180	0.0445	0.1166	0.1850	0.2408	0.5428	0.0774	0.0243	0.3798	-4.7522
Outside option	0.2496	0.0677	0.0128	0.0350	0.0658	0.0696	0.1701	0.0251	0.0053	0.1066	0.0279

Notes: The top panel gives matrix of price elasticities in the food at home segment for the set of products produced by the two firms that advertise most. The bottom panel gives matrix of price elasticities in the food on-the-go segment. Each cell contains the price elasticity of demand for the product indicated in column 1 with respect to the price of the product in row 1. Numbers are means across markets.

A.4.3 Mean marginal costs estimates

Table A.6: *Marginal costs*

	Price (£)	Cost (£)	Margin
Selected food at home products			
Walkers Regular:150-300g	1.11	0.25	0.77
		[0.22, 0.29]	[0.74, 0.80]
Walkers Regular:300g+	2.61	1.46	0.44
		[1.40, 1.52]	[0.42, 0.46]
Walkers Sensations:150-300g	1.26	-0.08	1.07
		[-0.13, -0.02]	[1.02, 1.11]
Walkers Sensations:300g+	2.79	1.05	0.62
		[0.97, 1.14]	[0.59, 0.65]
Walkers Doritos:150-300g	1.30	0.21	0.86
		[0.16, 0.25]	[0.83, 0.90]
Walkers Doritos:300g+	2.58	1.29	0.50
		[1.23, 1.34]	[0.48, 0.52]
Walkers Other:<150g	1.21	0.08	0.94
		[0.03, 0.12]	[0.91, 0.98]
Walkers Other:150-300g	2.50	1.15	0.54
		[1.09, 1.21]	[0.52, 0.57]
Walkers Other:300g+	1.24	0.03	0.98
		[-0.02, 0.08]	[0.94, 1.02]
Pringles:150-300g	1.77	0.45	0.75
		[0.40, 0.51]	[0.71, 0.78]
Pringles:300g+	3.17	1.59	0.50
		[1.52, 1.66]	[0.48, 0.52]
Food-on-the-go products			
Walkers Regular:34.5g	0.45	0.27	0.40
		[0.26, 0.28]	[0.38, 0.42]
Walkers Regular:50g	0.64	0.44	0.31
		[0.43, 0.45]	[0.29, 0.33]
Walkers Sensations:40g	0.62	0.42	0.33
		[0.41, 0.43]	[0.31, 0.35]
Walkers Doritos:40g	0.54	0.36	0.34
		[0.35, 0.37]	[0.32, 0.36]
Walkers Other:<30g	0.45	0.27	0.40
		[0.26, 0.28]	[0.37, 0.41]
Walkers Other:30g+	0.61	0.42	0.31
		[0.41, 0.43]	[0.29, 0.32]
KP:50g	0.52	0.37	0.28
		[0.37, 0.38]	[0.26, 0.29]
Golden Wonder:<40g	0.39	0.24	0.38
		[0.23, 0.25]	[0.35, 0.40]
Golden Wonder:40g+	0.73	0.54	0.26
		[0.52, 0.56]	[0.23, 0.28]
Other:<40g	0.50	0.35	0.30
		[0.34, 0.36]	[0.28, 0.31]
Other:40g+	0.66	0.50	0.24
		[0.49, 0.51]	[0.22, 0.25]

Notes: The top panel gives numbers for the food at home segment for the set of products produced by the two firms that advertise most. The bottom panel gives numbers for the food on-the-go segment. Margins are defined as $(p - mc)/p$. Numbers are means across markets. 95% confidence intervals are given in square brackets.

A.4.4 Profits

Table A.7 disaggregates the impact of the ban by firm and reports the average impact across months. The first panel reports pre ban numbers, showing mean advertising expenditures, the average price, total quantity of potato chips sold and total variable profits. The second panel details the percent change in quantity sold and variable profits resulting from the ban if firms do not re-optimize their prices in response. The final panel shows the impact on prices, quantity and variable profits following the ban in equilibrium, when firms are allowed to re-optimize prices.

When prices are held at their pre ban level the ban leads to a fall in the quantity sold and in variable profits of almost all firms in the market. Walkers is the exception - Walkers sees a relatively small decline in quantity sold, and its profits do not change. Holding prices fixed, Walkers, the largest and highest advertising firm, is not adversely affected by the ban. The reason for this is that other firms' advertising is strongly predatory towards Walkers products (particularly its largest brand, Walkers Regular). If Walkers unilaterally stopped advertising it would looss profits.[1] In the case of the ban, when firms reoptimize their prices, Walkers actually sees an increase in profitability, although other firms in the market are worse off than prior to the ban.

[1]For instance, if Walkers unilaterally set advertising expenditure of Walkers Regular to zero in a particular market holding prices fixed, they would see a fall in flow profits and a bigger fall in total discounted profits due the long lasting effect of advertising on demand.

Table A.7: *Advertising ban: Impact by firm*

	Walkers	Pringles	KP	Golden Wonder	Asda	Tesco	Other
Pre ban							
Advertising expenditure (£m)	1.01	0.45	0.21	0.01	0.00	0.01	0.16
Price (£)	1.76	1.86	1.32	1.38	1.39	1.27	1.49
	[1.76, 1.77]	[1.86, 1.86]	[1.31, 1.32]	[1.35, 1.41]	[1.39, 1.39]	[1.27, 1.27]	[1.49, 1.50]
Quantity (mKg)	7.33	0.94	2.72	0.26	0.39	0.76	2.51
	[7.16, 7.45]	[0.91, 0.98]	[2.65, 2.76]	[0.25, 0.28]	[0.37, 0.40]	[0.73, 0.79]	[2.45, 2.56]
Profits (£m)	30.15	2.43	6.83	0.70	0.74	1.46	6.85
	[28.46, 31.60]	[2.25, 2.62]	[6.52, 7.10]	[0.65, 0.75]	[0.70, 0.79]	[1.39, 1.53]	[6.48, 7.11]
Post ban: No firm response							
% change in quantity	-7.90	-42.70	-20.12	-24.92	-20.16	-20.85	-5.46
	[-12.71, -3.36]	[-46.86, -37.57]	[-24.73, -14.88]	[-29.47, -19.86]	[-25.97, -13.62]	[-26.52, -13.87]	[-10.99, 0.94]
% change in profits	-0.60	-25.31	-15.69	-28.42	-21.17	-21.49	-1.18
	[-5.92, 4.40]	[-31.31, -17.85]	[-20.62, -9.98]	[-33.27, -23.57]	[-26.76, -14.44]	[-27.26, -14.34]	[-6.80, 5.49]
Post ban: With firm response							
% change in price	-10.64	-17.81	-7.17	0.24	1.00	0.29	-5.97
	[-11.59, -6.70]	[-19.51, -16.11]	[-8.61, -6.00]	[-3.05, 2.88]	[0.33, 1.73]	[-0.43, 1.08]	[-7.19, -4.84]
% change in quantity	29.05	-19.96	-19.01	-33.85	-35.68	-34.93	-8.51
	[22.74, 35.13]	[-26.46, -11.31]	[-23.67, -13.54]	[-37.70, -29.08]	[-40.34, -30.23]	[-39.51, -29.35]	[-13.51, -2.77]
% change in profits	5.17	-30.86	-23.82	-32.26	-35.17	-35.05	-11.28
	[-0.04, 9.89]	[-36.33, -23.83]	[-28.59, -18.80]	[-36.89, -27.35]	[-39.72, -29.48]	[-39.52, -29.81]	[-16.57, -5.88]

Notes: "No firm response" refers to case of an advertising ban when prices are held at their pre ban level; "Firm response" refers to case of an advertising ban when firms reoptimize their prices. Price refers to the quantity weighted mean price set by the firm, quantity refers to the total amount of produce sold and profits are variable profits. Numbers are means across markets. 95% confidence intervals are given in square brackets.

B

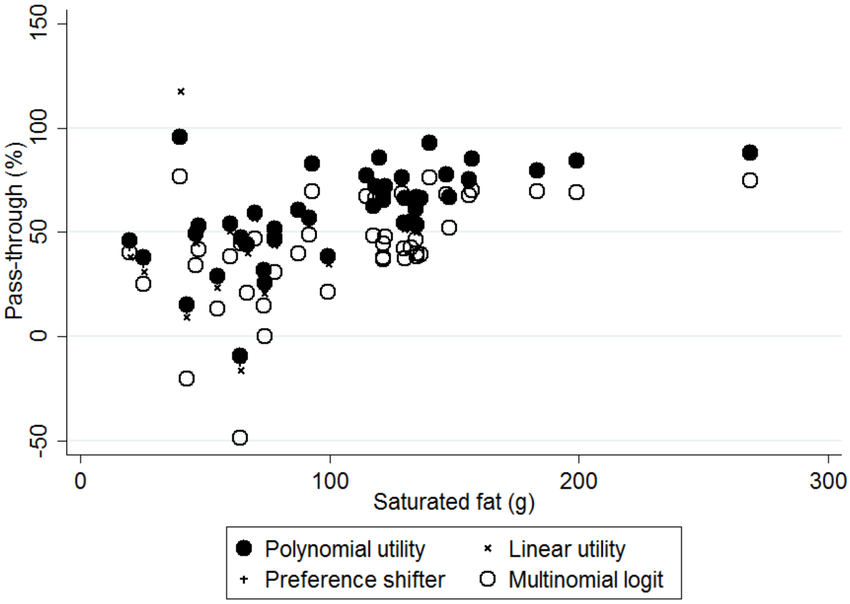
Appendix to "Income effects and the welfare consequences of tax in differentiated product oligopoly"

B.1 An ad valorem tax

In Section 3.4.7 we report results from simulating an excise tax that is proportional to products' saturated fat contents and that generates a 25% fall in purchases of saturated fat in the case of no firm pricing response (i.e. in the case of 100% pass-through). Here we report results results from simulating an ad valorem tax that is proportional to products' saturated fat content and that generates a 25% fall in purchases of saturated fat under 100% pass-through. Figure B.1 shows the patterns of pass-through across products for each model specification (the ad valorem tax analog of Figure 3.4). In contrast to the excise tax, the ad valorem tax is under-shifted to final prices - average pass-through under the polynomial utility model is 58%. However, as with the excise tax, the polynomial utility, linear utility and preference shifter models generate the same pattern of pass-through across products and the multinomial logit model generates pass-through that, on average, is lower (43% on average).

Table B.1 describes compensating variation from the ad valorem tax (the ad valorem tax analog of Table 3.5). As the ad valorem tax is under-shifted to consumer prices, compensating variation from the tax is less than for the excise tax. In common with the excise tax, the preference shifter and linear utility models fail to fully replicate how compensating variation varies across the expenditure distribution under the polynomial utility specification.

Figure B.1: *Ad valorem tax pass-through across products*



Notes: For each product in each market with compute the pass-through of the ad valorem tax. Figure is a scatter plot of products' mean pass-through across markets with their saturated fat contents.

Table B.1: *Compensating variation from ad valorem tax*

Specification	Mean compensating variation	Average deviation from mean compensating for quartile of expenditure distribution: 1	2	3	4
Polynomial utility	1.33	-0.59	-0.18	0.21	0.57
	[1.29, 1.39]	[-0.64, -0.56]	[-0.21, -0.15]	[0.18, 0.24]	[0.51, 0.64]
Linear utility	1.35	-0.19	-0.04	0.03	0.18
	[1.29, 1.43]	[-0.19, -0.16]	[-0.04, -0.03]	[0.03, 0.04]	[0.16, 0.20]
Preference shifter	1.34	-0.52	-0.25	0.06	0.65
	[1.27, 1.41]	[-0.56, -0.46]	[-0.26, -0.22]	[0.07, 0.09]	[0.61, 0.74]

Notes: Numbers give compensating variation for the average household associated with the simulated ad valorem tax. We measure expenditure as the households' mean weekly grocery expenditure. Numbers are for a calendar year. 95% confidence intervals are shown in brackets.

B.2 Additional tables

Here we include estimated market price elasticities and marginal costs for the polynomial utility, linear utility and preference shifter specifications. Table B.2 presents a matrix of average market own and cross price elasticities for the 10 products with the highest market share. It contains the matrix for each of the three model specifications. The numbers show: 1) Demand for all products is elastic, with own price elasticities ranging from -1.7 to -5.0. 2) Cross-price elasticities exhibit a high degree of variation, showing estimates are far from those from a conditional logit (in which there would be no within column variation in cross price elasticities). The cross-price elasticities also indicate a much higher degree of substitution within the butter products (Ar: Anchor NZ 500g, Ar: Lurpak spread ss 500g and Ar: Lurpak light ss 500g) than between them and the margarine products. 3) The three models yield similar estimates for market own and cross price elasticities. This contrasts with their predictions for household level elasticities which differ (see Table 3.4 and Figure 3 3.3).

Table B.3 presents the mean marginal cost estimates for the 10 largest market share products. They are based on the assumption that firms compete in a Nash-Bertrand game and therefore are a functions of the market level price elasticities and the ownership structure of products. Given the similarities in market elasticities between the three model specifications, it is not surprising that the models generate a similar set of marginal costs. Margins are estimated to be lower for the butter products than for the margarine products.

Table B.2: *Own and cross price elasticities*

Polynomial utility

	Ar: Anchor NZ 500g	Ar: Lurpak spread ss 500g	Ar: Lurpak light ss 500g	DC: Clover spread 500g	DC: Utterly Butterly 500g	Un: Flora buttery 500g	Un: Flora light 500g	Un: ICBINB 500g	Un: ICBINB Light 500g	Un: Stork 500g	Outside option
Ar: Anchor NZ 500g	-4.7723	0.1465	0.1441	0.0333	0.0357	0.0344	0.0340	0.0358	0.0356	0.0361	0.0364
Ar: Lurpak spread ss 500g	0.2470	-4.9456	0.2562	0.0570	0.0593	0.0600	0.0585	0.0605	0.0602	0.0603	0.0597
Ar: Lurpak light ss 500g	0.1757	0.1852	-5.0412	0.0412	0.0431	0.0436	0.0425	0.0439	0.0437	0.0437	0.0429
DC: Clover spread 500g	0.0346	0.0356	0.0357	-2.4253	0.0516	0.0524	0.0538	0.0493	0.0495	0.0505	0.0423
DC: Utterly Butterly 500g	0.0239	0.0236	0.0237	0.0331	-1.7729	0.0324	0.0331	0.0330	0.0328	0.0331	0.0304
Un: Flora buttery 500g	0.0412	0.0429	0.0431	0.0603	0.0576	-2.1381	0.0604	0.0598	0.0600	0.0584	0.0517
Un: Flora light 500g	0.0270	0.0279	0.0280	0.0411	0.0395	0.0403	-2.4811	0.0403	0.0404	0.0390	0.0320
Un: ICBINB 500g	0.0328	0.0332	0.0333	0.0443	0.0461	0.0460	0.0464	-1.8513	0.0477	0.0458	0.0417
Un: ICBINB light 500g	0.0166	0.0167	0.0168	0.0225	0.0233	0.0234	0.0236	0.0242	-1.9077	0.0232	0.0210
Un: Stork 500g	0.0131	0.0131	0.0131	0.0176	0.0180	0.0177	0.0176	0.0180	0.0180	-1.6520	0.0173

Linear utility

	Ar: Anchor NZ 500g	Ar: Lurpak spread ss 500g	Ar: Lurpak light ss 500g	DC: Clover spread 500g	DC: Utterly Butterly 500g	Un: Flora buttery 500g	Un: Flora light 500g	Un: ICBINB 500g	Un: ICBINB Light 500g	Un: Stork 500g	Outside option
Ar: Anchor NZ 500g	-4.7133	0.1452	0.1429	0.0335	0.0363	0.0351	0.0342	0.0366	0.0364	0.0369	0.0375
Ar: Lurpak spread ss 500g	0.2484	-4.8383	0.2586	0.0581	0.0611	0.0618	0.0597	0.0626	0.0623	0.0624	0.0622
Ar: Lurpak light ss 500g	0.1772	0.1873	-4.9295	0.0421	0.0445	0.0450	0.0435	0.0455	0.0453	0.0452	0.0448
DC: Clover spread 500g	0.0341	0.0350	0.0351	-2.3811	0.0518	0.0525	0.0536	0.0496	0.0498	0.0508	0.0423
DC: Utterly Butterly 500g	0.0235	0.0232	0.0232	0.0328	-1.7471	0.0322	0.0328	0.0330	0.0328	0.0332	0.0300
Un: Flora buttery 500g	0.0407	0.0423	0.0425	0.0599	0.0577	-2.0934	0.0601	0.0598	0.0601	0.0586	0.0511
Un: Flora light 500g	0.0267	0.0275	0.0277	0.0410	0.0397	0.0404	-2.4320	0.0405	0.0406	0.0392	0.0320
Un: ICBINB 500g	0.0324	0.0326	0.0327	0.0440	0.0461	0.0457	0.0460	-1.8185	0.0477	0.0459	0.0410
Un: ICBINB light 500g	0.0163	0.0165	0.0165	0.0224	0.0233	0.0232	0.0234	0.0241	-1.8736	0.0232	0.0206
Un: Stork 500g	0.0128	0.0128	0.0128	0.0173	0.0179	0.0175	0.0174	0.0178	0.0178	-1.6266	0.0169

Preference shifter

	Ar: Anchor NZ 500g	Ar: Lurpak spread ss 500g	Ar: Lurpak light ss 500g	DC: Clover spread 500g	DC: Utterly Butterly 500g	Un: Flora buttery 500g	Un: Flora light 500g	Un: ICBINB 500g	Un: ICBINB Light 500g	Un: Stork 500g	Outside option
Ar: Anchor NZ 500g	-4.7609	0.1458	0.1434	0.0332	0.0357	0.0344	0.0339	0.0358	0.0356	0.0361	0.0367
Ar: Lurpak spread ss 500g	0.2466	-4.9109	0.2570	0.0570	0.0594	0.0600	0.0586	0.0606	0.0604	0.0604	0.0602
Ar: Lurpak light ss 500g	0.1756	0.1858	-5.0039	0.0413	0.0432	0.0437	0.0426	0.0440	0.0438	0.0438	0.0433
DC: Clover spread 500g	0.0347	0.0356	0.0357	-2.4295	0.0519	0.0526	0.0540	0.0495	0.0497	0.0508	0.0426
DC: Utterly Butterly 500g	0.0239	0.0237	0.0237	0.0333	-1.7807	0.0326	0.0333	0.0332	0.0331	0.0334	0.0307
Un: Flora buttery 500g	0.0413	0.0430	0.0432	0.0606	0.0580	-2.1433	0.0608	0.0602	0.0604	0.0588	0.0522
Un: Flora light 500g	0.0271	0.0279	0.0281	0.0412	0.0397	0.0405	-2.4840	0.0405	0.0406	0.0392	0.0323
Un: ICBINB 500g	0.0330	0.0333	0.0333	0.0446	0.0464	0.0463	0.0467	-1.8579	0.0481	0.0462	0.0421
Un: ICBINB light 500g	0.0166	0.0168	0.0168	0.0227	0.0234	0.0235	0.0237	0.0244	-1.9143	0.0234	0.0211
Un: Stork 500g	0.0131	0.0132	0.0132	0.0177	0.0182	0.0179	0.0178	0.0182	0.0182	-1.6600	0.0174

Notes: Each cell contains the price elasticity of demand for the product indicated in row 1 with respect to the price of the product in column 1. Numbers are means across markets.

Table B.3: *Marginal costs: Top 10 market share products*

	Price	Polynomial utility Cost	Polynomial utility Margin	Linear utility Cost	Linear utility Margin	Preference shifter Cost	Preference shifter Margin
Ar: Anchor NZ 500g	1.99	1.49	0.25	1.49	0.26	1.49	0.25
Ar: Lurpak spread ss 500g	2.15	1.64	0.24	1.63	0.24	1.64	0.24
Ar: Lurpak light ss 500g	2.17	1.66	0.24	1.64	0.24	1.65	0.24
DC: Clover spread 500g	1.20	0.67	0.44	0.66	0.45	0.67	0.44
DC: Utterly Butterly 500g	0.80	0.32	0.60	0.31	0.61	0.32	0.60
Un: Flora buttery 500g	1.02	0.46	0.55	0.44	0.57	0.46	0.55
Un: Flora light 500g	1.22	0.63	0.48	0.62	0.50	0.63	0.48
Un: ICBINB 500g	0.85	0.31	0.64	0.30	0.65	0.31	0.63
Un: ICBINB light 500g	0.87	0.33	0.62	0.31	0.64	0.33	0.62
Un: Stork 500g	0.72	0.20	0.72	0.19	0.74	0.20	0.72

Notes: Margins are defined as $(p - mc)/p$. Numbers are market share weighted means.

C

Appendix to "Partial consumption smoothing over the Great Recession"

C.1 Additional tables

Table C.1: *Changes in the Healthy Eating Index*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
				Household type				Employment status and income			
	Max score	Mean score	All households	Pre-school children	School age children	No children, no pensioners	Pensioner households	Working; income: High	Middle	Low	Unemp.
Mean HEI in 2005-2007	100	49.0	49.0	48.7	46.1	47.8	51.5	49.6	48.0	46.6	46.7
Total change to 2010-2012			0.72	1.52	1.03	1.46	-0.23	0.87	1.03	2.01	1.11
of which											
'Good' change			1.45	3.02	1.90	1.93	0.91	1.78	1.78	2.44	1.67
'Bad' change			-0.72	-1.51	-0.87	-0.46	-1.14	-0.91	-0.75	-0.43	-0.56
			which consists of changes in the component scores:								
Total fruit	5	3.06	-0.02	0.12	-0.05	0.02	-0.05	-0.05	-0.05	-0.07	0.04
Whole fruit	5	3.36	0.08	0.26	-0.07	0.18	0.03	0.11	0.05	0.03	0.19
Total vegetables	5	3.20	-0.13	-0.34	-0.05	-0.04	-0.20	-0.12	-0.12	-0.13	-0.02
Dark green/orange veg	5	1.61	0.00	-0.07	0.07	0.09	-0.09	0.05	-0.00	0.09	0.10
Total grains	5	3.69	-0.03	0.08	0.01	-0.04	-0.07	-0.07	-0.00	0.10	0.02
Whole grains	5	1.55	-0.11	-0.38	-0.14	-0.08	-0.06	-0.26	-0.18	-0.03	-0.08
Milk	10	5.28	-0.05	-0.59	-0.31	0.07	0.06	-0.21	-0.15	-0.11	-0.12
Meat	10	7.96	-0.22	-0.13	-0.06	-0.17	-0.33	-0.20	-0.13	-0.03	-0.17
Oils	10	4.93	-0.18	0.09	-0.20	-0.14	-0.30	0.02	-0.11	-0.06	-0.17
Sodium	10	6.42	0.93	1.31	0.93	1.00	0.77	1.10	1.11	0.95	0.93
Saturated fat	10	2.70	0.27	0.80	0.60	0.24	0.06	0.34	0.41	0.66	0.38
Calories from SoFAAS	20	5.22	0.18	0.36	0.28	0.33	-0.05	0.16	0.21	0.61	0.43

Notes: Row 1 shows the mean overall HEI score for all households (column (3)) and each household type (columns (4)-(11)) in 2005-2007; row 2 shows the average within household change in the HEI from 2005-2007 to 2010-2012. This is the sum of the changes in the component scores; these are shown in the bottom panel of the table. 'Good change' (shown in row 3) is the sum of the positive changes in the bottom panel; 'Bad change' (shown in row 4) is the sum of the negative changes in the bottom panel. Column (1) shows the maximum score for each component; these sum to 100 (the maximum score for the HEI). Column (2) shows the mean component score in 2005-2007 across all households. "Calories from SoFAAS" is the share of calories from solid fat, added sugar and alcohol. The group "Pensioners" within the "Employment status and income" division are identical to the group of households in "Pensioner households", shown in column (7).

Table C.2: *Contribution of choice variables to change in price paid per calorie, by household composition*

	Households with children		Households without children	
	Youngest child is: Pre-school	School age	No pensioners	Pensioners
Shopping effort:				
Number of shopping trips	-0.13	-0.14	0.02	-0.01
Number of chains visited	-0.02	-0.02	0.04	0.03
Savings from discounter	-0.13	-0.02	-0.07	-0.12
Savings from sales	-1.38	-1.18	-0.99	-0.78
Total	-1.66	-1.36	-1.00	-0.87
Nutrient characteristics:				
Protein	-1.42	1.04	-0.39	-0.77
Saturated fat	-0.71	-0.49	-0.20	-0.07
Unsaturated fat	-0.01	0.12	0.01	0.04
Sugar	0.10	-0.04	0.03	0.00
Fibre	-0.54	-0.36	-0.42	-0.32
Salt	0.08	0.07	0.07	0.04
Fruit	0.33	-0.07	0.46	0.30
Vegetables	-0.44	-0.06	-0.14	-0.34
Dairy	0.00	0.01	-0.00	-0.00
Cheese and fats	0.06	-0.01	-0.01	0.02
Poultry and fish	-0.13	-0.24	-0.09	-0.09
Red meat and nuts	0.02	0.01	0.03	0.06
Drinks	0.06	0.30	-0.13	-0.09
Prepared sweet	0.31	-0.09	0.11	0.18
Prepared savory	0.11	0.06	-0.27	0.31
Alcohol	-0.58	0.10	-0.05	-0.19
Total	-2.76	0.37	-0.99	-0.93
Other characteristics:				
Share from generic products	-1.11	-0.43	-0.75	-1.02
Share of calories from big packs	0.34	0.19	0.17	0.29
Total	-0.77	-0.24	-0.59	-0.73
Total	-5.19	-1.23	-2.57	-2.53

Notes: The table reports the contribution each variable made to the fall in price paid per calorie. The contribution is given by the product of the coefficient in column 3 of Table 5.1 and average change in log of the transformed variable, controlling for fixed effects (multiplied by 100). "Pre-school" denotes children aged between 0 and 5; "school age" between 6 and 17. "Pensioner" households are those in which at least one member is aged 65 or over.

Table C.3: *Contribution of choice variables to change in price paid per calorie, by employment status and income*

	Working; income:			
	High	Middle	Low	Unemployed
Shopping effort:				
Number of shopping trips	-0.02	-0.05	-0.03	-0.07
Number of chains visited	-0.00	0.02	0.04	0.01
Savings from discounter	-0.10	-0.07	-0.09	-0.04
Savings from sales	-1.14	-1.12	-1.06	-1.05
Total	-1.26	-1.22	-1.14	-1.15
Nutrient characteristics:				
Protein	-0.64	-0.18	0.69	-0.09
Saturated fat	-0.24	-0.28	-0.53	-0.37
Unsaturated fat	0.17	0.05	0.04	0.02
Sugar	0.00	0.02	-0.01	0.01
Fibre	-0.42	-0.39	-0.35	-0.48
Salt	0.08	0.08	0.06	0.05
Fruit	-0.04	0.13	0.19	0.35
Vegetables	-0.21	-0.20	-0.21	-0.08
Dairy	0.00	0.00	0.00	0.00
Cheese and fats	-0.10	-0.03	0.01	0.03
Poultry and fish	-0.13	-0.16	-0.22	-0.08
Red meat and nuts	0.02	0.03	0.01	0.02
Drinks	0.08	-0.02	0.07	-0.05
Prepared sweet	0.09	0.09	-0.05	0.07
Prepared savoury	-0.24	-0.23	-0.13	-0.03
Alcohol	0.04	0.31	0.22	-0.40
Total	-1.56	-0.79	-0.23	-1.02
Other:				
Share from generic products	-0.88	-0.72	-0.99	-0.37
Share of groceries from big pack sizes	0.12	0.16	0.31	0.27
Total	-0.76	-0.57	-0.68	-0.10
Total	-3.58	-2.57	-2.04	-2.27

Notes: The table reports the contribution each variable made to the fall in price paid per calorie. The contribution is given by the product of the coefficient in column 3 of Table 5.1 and average change in log of the transformed variable, controlling for fixed effects (multiplied by 100). Working households are those in which the head of the household works more than 8 hours a week. Income is measure using social grade; grade AB/C/DE correspond to high/middle/low income.

Bibliography

Abramovsky, L., R. Griffith, G. Macartney, and H. Miller (2008). Globalisation and innovation. *IFS Working Paper, W08/10*.

Ackerberg, D. (2001). Empirically Distinguishing Informative and Prestige Effects of Advertising. *Rand Journal of Economics 32*(2), 316 – 333.

Ackerberg, D. (2003). Advertising, Learning, and Consumer Choice in Experience Good Markets: An Empirical Examination. *International Economic Review 44*(3), 1007 – 1040.

Adda, J., H.-M. Gaudecker, and J. Banks (2009). The impact of income shocks on health: evidence from cohort data. *Journal of the European Economic Association 7*(6), 1361–1399.

Aguiar, M. and E. Hurst (2005). Consumption versus expenditure. *Journal of Political Economy 113*(5), 919–948.

Aguiar, M. and E. Hurst (2007). Life-cycle prices and production. *American Economic Review*, 1533–1559.

Aguiar, M., E. Hurst, and L. Karabarbounis (2013). Time use during the great recession. *American Economic Review 103*(5), 1664–1696.

Anderson, S., F. Ciliberto, J. Liaukonyte, and R. Renault (2012). Push-Me Pull-You: Comparative Advertising in the OTC Analgesics Industry. *CEPR Discussion Paper 8988*.

Anderson, S. P., A. De Palma, and B. Kreider (2001). Tax incidence in differentiated product oligopoly. *Journal of Public Economics 81*(2), 173–192.

Anderson, S. P. and R. Renault (2006). Advertising content. *American Economic Review 96*(1), 93–113.

Arambepola, C., M. Scarborough, and M. Rayner (2008). Validating a nutrient profile model. *Public Heath Nutrition 11*, 371–378.

Bagwell, K. (2007). The economic analysis of advertising. *Handbook of Industrial Organization 3*, 1701–1844.

Bajari, P. and C. L. Benkard (2005). Demand estimation with heterogeneous consumers and unobserved product characteristics: A hedonic approach. *Journal of Political Economy 113*(6), 1239–1276.

Banks, J., R. Blundell, and A. Lewbel (1997). Quadratic Engel Curves and Consumer Demand. *Review of Economics and Statistics 79*(4), 527–539.

Beatty, T. K. M., B.-H. Lin, and T. A. Smith (2014). Is diet quality improving? distributional changes in the united states, 1989-2008. *American Journal of Agricultural Economics*.

Becker, G. S. (1965). A theory of the allocation of time. *Economic Journal 75*, 493–517.

Becker, J. and C. Fuest (2007). Quality versus quantity: the composition effect of corporate taxation on foreign direct investment. Technical report, CESifo Working Paper Series 2126.

Becker, G., S. and K. M. Murphy (1993). A Simple Theory of Advertising as a Good or Ban. *The Quarterly Journal of Economics 108*(4), 941–964.

Bernard, A. B., J. B. Jensen, S. J. Redding, and P. K. Schott (2007a). Comparative advantage and heterogeneous firms. *The Review of Economic Studies 74*(1), 31–66.

Bernard, A. B., J. B. Jensen, S. J. Redding, and P. K. Schott (2007b). Firms in International Trade. *Journal of Economic Perspectives 21*(3), 105–130.

Bernheim, B. and A. Rangel (2004). Addication and cue-triggered decision processes. *American Economic Review 94*, 1558?1590.

Bernheim, B. and A. Rangel (2005). Behavioral public economics welfare and policy analysis with non-standard decision-makers. *NBER Working Paper 11518*.

Bernheim, B. and A. Rangel (2009). Beyond Revealed Preference: Choice-Theoretic Foundatations for Behavioral Welfare Economics. *Quarter Journal of Economics 124*(1), 51–104.

Bernheim, B. D., J. Skinner, and S. Weinberg (2001). What accounts for the variation in retirement wealth among US households? *American Economic Review*, 832–857.

Berry, S. (1994). Estimating Discrete-Choice Models of Product Differentiation. *RAND Journal of Economics 25*(2), 242–262.

Berry, S. and P. Haile (2010). Nonparametric Identification of Multinomial Choice Demand Models with Heterogeneous Consumers. *Cowles Foundation Discussion Paper 1718*.

Berry, S. and P. Haile (2014). Identification in Differentiated Products Markets Using Market Level Data. *Econometrica 82*(5), 1749–1797.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile Prices in Market Equilibrium. *Econometrica 63*(4), 841–890.

Berry, S., J. Levinsohn, and A. Pakes (2004). Differentiated Products Demand Systems from Combination of Micro and Macro Data: The New Car Market. *Journal of Political Economy 112*(1), 68–105.

Besley, T. and H. Rosen (1999). Sales taxes and prices: an empirical analysis. *National tax journal 52*(2), 157–178.

Bhattacharya, D. (2015). Nonparametric Welfare Analysis for Discrete Choice. *Econometrica 83*(2), 617–649.

Blundell, R., R. Griffith, and J. Van Reenen (1999). Market share, market value and innovation in a panel of British manufacturing firms. *The Review of Economic Studies 66*(3), 529–554.

Blundell, R., L. Pistaferri, and I. Preston (2008). Consumption inequality and partial insurance. *American Economic Review 98*(5), 1887–1921.

Blundell, R., L. Pistaferri, and I. Saporta-Eksten (2014). Consumption inequality and family labor supply. In *2014 Meeting Papers*, Number 322. Society for Economic Dynamics.

Blundell, R. and J. Powell (2004). Endogeneity in Semiparametric Binary Response Models. *Review of Economic Studies 71*, 655–679.

Blundell, R. and I. Preston (1998). Consumption inequality and income uncertainty. *Quarterly Journal of Economics 113*(2), 603–640.

Bohm, T, K. T. and N. Riedel (2012). The Impact of Corporate Taxes on R&D and Patent Holdings. *mimeo*.

Braithwaite, D. (1928). The economic effects of advertisement. *Economic Journal 38*, 16–37.

Brewer, M., J. Browne, A. Hood, R. Joyce, and L. Sibieta (2013). The Short-and Medium-Term Impacts of the Recession on the UK Income Distribution. *Fiscal Studies 34*(2), 179–201.

Bucovetsky, S. and A. Haufler (2007). Preferential Tax Regimes with Asymmetric Countries. *National Tax Journal*, 789–795.

Cairns, G., K. Angus, and G. Hastings (2009). *The extent, nature and effects of food promotion to children: a review of the evidence to December 2008*. World Health Organization: Geneva.

Case, A., A. Fertig, and C. Paxson (2005). The lasting impact of childhood health and circumstance. *Journal of Health Economics 24*(2), 365–389.

Chamberlin, E. (1933). *The Theory of Monopolistic Competition.* Harvard University Press: Cambridge, MA.

Cohen, M. (2012). Corporate governance law: firm heterogeneity and the market for corporate domicile. *SSRN Working Paper 2116536*.

Coibion, O., Y. Gorodnichenko, and G. H. Hong (2014). The cyclicality of sales, regular and effective prices: Business cycle and policy implications. *American Economic Review (forthcoming)*.

Competition Commission (2000). Supermarkets: A Report on the Supply of Groceries from Multiple Stores in the United Kingdom.

Crossley, T. F., H. Low, and C. O'Dea (2013). Household consumption through recent recessions. *Fiscal Studies 34*(2), 203–229.

Currie, J. (2009). Healthy, wealthy, and wise: Is there a causal relationship between child health and human capital development? *Journal of Economic Literature 47*(1), 87–122.

Dagsvik, J. K. and A. Karlström (2005). Compensating variation and Hicksian choice probabilities in random utility models that are nonlinear in income. *The Review of Economic Studies 72*(1), 57–76.

De Mooij, R. A. and S. Ederveen (2008). Corporate tax elasticities: a reader's guide to empirical findings. *Oxford Review of Economic Policy 24*(4), 680–697.

Deaton, A. and J. Muellbauer (1980). An almost ideal demand system. *American Economic Review*, 312–326.

Dehejia, R. and A. Lleras-Muney (2004). Booms, busts, and babies' health. *Quarterly Journal of Economics 119*(3), 1091–1130.

Delipalla, S. and M. Keen (1992). The comparison between ad valorem and specific taxation under imperfect competition. *Journal of Public Economics 49*(3), 351–367.

Delipalla, S. and O. O'Donnell (2001). Estimating tax incidence, market power and market conduct: the European cigarette industry. *International Journal of Industrial Organization 19*(6), 885–908.

Department of Health (1991). Dietary reference values for food energy and nutrients for the united kingdom. *Report on Health and Social Subjects 25.*

Desai, M. A., C. F. Foley, and J. R. Hines Jr (2006). The demand for tax haven operations. *Journal of Public Economics 90*(3), 513–531.

Devereux, M. P. and G. Maffini (2006). The Impact of Taxation on the Location of Capital, Firms and Profit: A Survey of Empirical Evidence. *Oxford University Centre for Business Taxation Working Paper Series WP 07/02*.

Dischinger, M. and N. Riedel (2011). Corporate taxes and the location of intangible assets within multinational firms. *Journal of Public Economics 95*(7), 691–707.

Dixit, A. (1980). The Role of Investment in Entry Deterrence. *Economic Journal 90*, 95–106.

Dixit, A. and V. Norman (1978). Advertising and Welfare. *Bell Journal of Economics 9*(1), 1–17.

Doraszelski, U. and S. Markovich (2007). Advertising dynamics and competitive advantage. *RAND Journal of Economics 38*(3), 557–592.

Doraszelski, U. and M. Satterthwaite (2003). Foundations of Markov-Perfect Industry Dynamics: Existence, Purification, and Multiplicity. *Working Paper, Stanford, Calif.: Hoover Institution*.

Doraszelski, U. and M. Satterthwaite (2010). Computable Markov-perfect industry dynamics. *Rand Journal of Economics 41*(2), 215–243.

Dubé, J., G. Hitsch, and P. Manchanda (2005). An empirical model of advertising dynamics. *Quantitative Marketing and Economics 3*, 107–144.

Dubois, P., R. Griffith, and A. Nevo (2014). Do Prices and Attributes Explain International Differences in Food Purchases? *American Economic Review 104*(3), 832–67.

Eckard, W. (1991). Competition and the Cigarette TV Advertising Ban. *Economic Inquiry 29*, 119–133.

Eden, B. (2013). Price dispersion and demand uncertainty: Evidence from US scanner data. *Vanderbilt University Department of Economics Working Paper 13-00015.*

Erdem, T., M. Keane, and B. Sun (2008). The impact of advertising on consumer price sensitivity in experience goods markets. *Quantitative Marketing and Economics 6*(2), 139–176.

Ericson, R. and A. Pakes (1995). Markov-perfect industry dynamics: A framework for empirical work. *Review of Economic Studies 62*, 53–82.

Ernst, C. and C. Spengel (2011). Taxation, R&D tax incentives and patent application in Europe. *ZEW Discussion Paper* (11-024).

Evers, L., H. Miller, and C. Spengel (2013). Intellectual property box regimes: Effective tax rates and tax policy considerations. *ZEW Discussion Paper* (13-07).

Fabinger, M. and E. G. Weyl (2014). A tractable approach to pass-through patterns with applications to international trade. *mimeo*.

Fabra, N. and M. Reguant (2014). Pass-Through of Emissions Costs in Electricity Markets. *American Economic Review*, Forthcoming.

Finnerty, C, M. P. P. M. and R. Russo (2007). Fundamentals of International Tax Planning. IBFD.

Friedman, J. (1983). Advertising and Oligopolistic Equilibrium. *Bell Journal of Economics 14*, 461–473.

Fudenberg and Tirole (1984). The Fat-cat Effect, The Puppy-dog Ploy, and The Lean and Hungry Look. *American Economic Review Papers and Proceedings 74*, 361–366.

Gabaix, X. and D. Laibson (2006). Shrouded attributes, consumer myopia, and information suppression in competitive markets. *Quarterly Journal of Economics 121*(2), 505–540.

Gicheva, D., J. Hastings, and S. Villas-Boas (2010). Investigating income effects in scanner data: Do gasoline prices affect grocery purchases? *American Economic Review Papers & Proceedings 100*(2), 480–484.

Ginarte, J. C. and W. G. Park (1997). Determinants of patent rights: A cross-national study. *Research policy 26*(3), 283–301.

Glaeser, E. L. and G. Ujhelyi (2010). Regulating misinformation. *Journal of Public Economics 94*, 247 – 257.

Goldberg, P. K. and R. Hellerstein (2008). A structural approach to explaining incomplete exchange-rate pass-through and pricing-to-market. *American Economic Review*, 423–429.

Goldberg, P. K. and R. Hellerstein (2013). A structural approach to identifying the sources of local currency price stability. *The Review of Economic Studies 80*(1), 175–210.

Goldberg, P. K. and F. Verboven (2001). The evolution of price dispersion in the European car market. *The Review of Economic Studies 68*(4), 811–848.

Goris, J., S. Petersen, E. Stamatakis, and J. Veerman (2010). Television food advertising and the prevalence of childhood overweight and obesity: a multi country comparison. *Public Health Nutrition 13*, 1003–12.

Gorman, W. M. (1980). A Possible Procedure for Analysing Quality Differentials in the Egg Market. *The Review of Economic Studies 47*(5), 843–856.

Gortmaker, S., B. Swinburn, D. Levy, R. Carter, P. Mabry, D. Finegood, T. Huang, T. Marsh, and M. Moodie (2011). Changing the future of obesity: science, policy and action. *Lancet 378*, 838–847.

Graham, J. R. and A. L. Tucker (2006). Tax shelters and corporate debt policy. *Journal of Financial Economics 81*(3), 563–594.

Griffith, R., M. Krol, and K. Smith (2014). Store brand penetration: the role of advertising. *IFS Mimeo*.

Griffith, R., E. Leibtag, A. Leicester, and A. Nevo (2009). Consumer shopping behavior: how much do consumers save? *Journal of Economic Perspectives 23*(2), 99–120.

Griffith, R. and M. O'Connell (2009). The Use of Scanner Data for Research into Nutrition. *Fiscal Studies 30*, 339–365.

Grubert, H. (2003). Intangible income, intercompany transactions, income shifting, and the choice of location. *National Tax Journal*, 221–242.

Grubert, H. and R. Altshuler (2008). Corporate Taxes in the world economy: reforming the taxation of cross-border income. *Departmental Working Papers 200626, Rutgers University, Department of Economics*.

Grubert, H. and J. Mutti (1991). Taxes, tariffs and transfer pricing in multinational corporate decision making. *The Review of Economics and Statistics*, 285–293.

Hall, B. H., G. Thoma, and S. Torrisi (2007). The Market Value of Patents and R&D: Evidence from European Firms. *NBER Working Paper No. 13426*.

Hall, R. E. and D. W. Jorgenson (1967). Tax policy and investment behavior. *American Economic Review*, 391–414.

Hall, R. E. and F. S. Mishkin (1982). The sensitivity of consumption to transitory income: Estimates from panel data on households. *Econometrica 50*(2), 461–81.

Hausman, J. (1996). *Valuation of new goods under perfect and imperfect competition.*

Hausman, J. and W. Newey (2014). Individual heterogeneity and average welfare. *Cemmap working paper*.

Herriges, J. A. and C. L. Kling (1999). Nonlinear income effects in random utility models. *Review of Economics and Statistics 81*(1), 62–72.

Hines, J. R. (1996). Tax policy and the activities of multinational corporations. Technical report.

Hines, J. R. (1999). Lessons from behavioral responses to international taxation. *National Tax Journal*, 305–322.

Hines, J. R. and A. B. Jaffe (2000). International taxation and the location of inventive activity. In *International Taxation and Multinational Activity*, pp. 201–230. University of Chicago Press.

Hines, J. R. and E. M. Rice (1994). Fiscal Paradise: Foreign Tax Havens and American Business. *The Quarterly Journal of Economics 109*(1), 149–82.

Jappelli, T. and L. Pistaferri (2010). The consumption response to income changes. *NBER Working Paper 15739*.

Kahneman, D., P. P. Wakker, and R. Sarin (1997). Back to Bentham? Explorations of experienced utility. *Quarterly Journal of Economics*, 375–405.

Kaldor, N. (1950). The economic aspects of advertising. *Review of Economic Studies 18*(1), 1–27.

Kaplan, G. and G. Menzio (2014a). The morphology of price dispersion. *NBER Working Paper 19877*.

Kaplan, G. and G. Menzio (2014b). Shopping externalities and self-fulfilling unemployment fluctuations. *Journal of Political Economy (forthcoming)*.

Karkinsky, T. and N. Riedel (2012). Corporate taxation and the choice of patent location within multinational firms. *Journal of International Economics 88*(1), 176–185.

Kenkel, D. S. (2005). Are alcohol tax hikes fully passed through to prices? Evidence from Alaska. *American Economic Review*, 273–277.

Kim, D. and R. W. Cotterill (2008). Cost Pass-Through in Differentiated Product Markets: The Case of US Processed Cheese. *Journal of Industrial Economics 56*(1), 32–48.

Krautheim, S. and T. Schmidt-Eisenlohr (2011). Heterogeneous firms,profit shifting FDI and international tax competition. *Journal of Public Economics 95*(1), 122–133.

Kuchler, F. (2011). Is it food quality or quantity that responds to changing income? *Applied Economic Perspectives and Policy 33*(2), 205–221.

Kumcu, A. and P. Kaufman (2011). Food spending adjustments during recessionary times. *Amber Waves 9*(3), 10–17.

Lancaster, K. (1971). *Consumer demand: A new approach.* Columbia University Press New York.

Lanjouw, J. O. and M. Schankerman (2004). Patent quality and research productivity: Measuring innovation with multiple indicators*. *The Economic Journal 114*(495), 441–465.

Leicester, A. and Z. Oldfield (2009). An analysis of consumer panel data. *IFS Working Papers W09/09*.

Lewbel, A. and K. Pendakur (2009). Tricks with Hicks: The EASI Demand System. *American Economic Review 99*(3), pp. 827–863.

Lipsey, R. E. (2010). Measuring the location of production in a world of intangible productive assets, FDI, and intrafirm trade. *Review of Income and Wealth 56*(s1), S99–S110.

Lock, K., D. Stuckler, K. Charlesworth, and M. McKee (2009). Potential causes and health effects of rising global food prices. *British Medical Journal 339.*

Marshall, A. (1921). *Industry and Trade: A study of Industrial Technique and Business Organization and of Their Influences on the Conditions of Various Classes and Nations.* MacMillan and Co.: London.

Maskin, E. and J. Tirole (1988). A theory of dynamic oligopoly, I: Overview and quantity competition with large fixed costs. *Econometrica 56*, 549–569.

Maskin, E. and J. Tirole (2001). Markov Perfect Equilibrium I: Observable Actions. *Journal of Economic Theory 100*(2), 191–219.

McClure, S., D. Laibson, G. Loewenstein, and J. Cohen (2004). Separate neural systems value immediate and delayed monetary rewards. *Science 306*, 503–507.

McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behavior. In P. Zarembka (Ed.), *Frontier in Econometrics*, pp. 198–272. New York: Academic Press.

McFadden, D. (1981). Econometric models of probabilistic choice among products. In C. F. Manski and D. McFadden (Eds.), *Structural analysis of discrete data with econometric applications*, pp. 198–272. Cambridge: MIT Press.

McFadden, D. (1999). Computing willingness-to-pay in random utility models. *Trade, Theory, and Econometrics: Essays in Honour of John S. Chipman 15*, 253.

McFadden, D. and K. Train (2000). Mixed MNL models for discrete response. *Journal of applied Econometrics 15*(5), 447–470.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica 71*(6), 1695–1725.

Milyo, J. and J. Waldfogel (1999). The Effect of Price Advertising on Prices: Evidence in the Wake of 44 Liquormart. *American Economic Review 89*, 1081–96.

Motta, M. (2007). Advertising bans. *UPF Working Papers No 205*.

Mullainathan, S., J. Schwartzstein, and W. J. Congdon (2012). A Reduced Form Approach to Behavioral Public Finance. *Annual Review of Economics 4*(17), 1–30.

Mutti, J. and H. Grubert (2009). The effect of taxes on royalties and the migration of intangible assets abroad. In *International trade in services and intangibles in the era of globalization*, pp. 111–137. University of Chicago Press.

Nakamura, E. and D. Zerom (2010). Accounting for incomplete pass-through. *The Review of Economic Studies 77*(3), 1192–1230.

National Academies, T. (2006). Committee on Food Marketing and the Diets of Children and Youth report on Food marketing to children and youth: threat or opportunity?

Nelson, P. (1995). Information and consumer behavior. *Journal of Political Economy 78*, 311–329.

Nevo, A. (2000). Mergers with differentiated products: The case of the ready-to-eat cereal industry. *The RAND Journal of Economics*, 395–421.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica 69*(2), 307–342.

Nevo, A. and A. Wong (2014). The Elasticity of Substitution Between Time and Market Goods: Evidence from the Great Recession. *memeo*.

OECD (2006). Report on the Attribution of Profits to Permanent Establishments, Parts I (General Considerations), II (Banks) and III (Global Trading). Technical report, OECD Publishing.

OECD (2012). OECD Factbook 2011-2012: Economic, Environmental and Social Statistics. Technical report, OECD Publishing.

Ofcom (2007). *Television advertising of food and drink products to children: Final statement.* London: Ofcom.

Pakes, A. (1985). On Patents, R & D, and the Stock Market Rate of Return. *Journal of Political Economy*, 390–409.

Pakes, A. (1986). Patents Options: Some Estimates of the Value of Holding European Patents Stocks. *Econometrica 54*(4), 755–784.

Park, W. G. (2008). International patent protection: 1960–2005. *Research policy 37*(4), 761–766.

Petrin, A. (2002). Quantifying the Benefits of New Products: The Case of the Minivan. *Journal of Political Economy 110*(4), 705–729.

Petrin, A. and K. Train (2010). A Control Function Approach to Endogeneity in Consumer Choice Models. *Journal of Marketing Research XLVII*, 3–13.

Qi, S. (2013). The impact of advertising regulation on industry: The cigarette advertising ban of 1971. *The RAND Journal of Economics 44*(2), 215–248.

Rayner, M., P. Scarborough, A. Boxer, and L. Stockley (2005). Nutrient profiles: Development of final model. Technical report, Food Standards Agency: London.

Rayner, M., P. Scarborough, and T. Lobstein (2009). The UK Ofcom Nutrient Profiling Model. Technical report, Nuffield Department of Population Healthy, Oxford University.

Robinson, J. (1933). *Economics of Imperfect Competition.* MacMillan and Co., London.

Rojas, C. and E. Peterson (2008). Demand for Differentiated Products: Price and Advertising Evidence from the U.S. beer Market. *International Journal of Industrial Organization 26*, 288–307.

Ruhm, C. J. (2000). Are recessions good for your health? *Quarterly Journal of Economics 115*(2), 617–650.

Ryan, S. (2012). The Costs of Environmental Regulation in a Concentrated Industry. *Econometrica 80*, 1019–1062.

Sass, T. and D. Saurman (1995). Advertising Restrictions and Concentration: The Case of Malt Beverages. *Review of Economics and Statistics 77*, 66–81.

Schmalensee, R. (1983). Advertising and Entry Deterrence: An Exploratory Model. *Journal of Political Economy 91*, 636–53.

Seade, J. (1985). Profitable cost increases and the shifting of taxation: equilibrium response of markets in oligopoly. Technical report, University of Warwick, Department of Economics.

Small, K. A. and H. S. Rosen (1981). Applied welfare economics with discrete choice models. *Econometrica 49*(1), 105–130.

Sovinsky-Goeree, M. (2008). Limited Information and Advertising in the US Personal Computer Industry. *Econometrica 76*(5), 1017–1074.

Spiegler, R. (2006). Competition over agents with bounded rationality. *Theoretical Economics I*, 207–231.

Stanley, L. R. and J. Tschirhart (1991). Hedonic prices for a nondurable good: The case of breakfast cereals. *Review of Economics and Statistics 73*(3), 537–41.

Stigler, G. and G. Becker (1977). De gustibus non est disputandum. *American Economic Review 67*, 76–90.

Stigler, G. J. (1961). The economics of information. *Journal of Political Economy 69*(3), 213–225.

Strauss, J. and D. Thomas (1998). Health, nutrition, and economic development. *Journal of Economic Literature*, 766–817.

Sweeting, A. (2013). Dynamic Product Positioning in Differentiated Product Markets: The Effect of Fees for Musical Performance Rights on the Commercial Radio Industry. *Econometrica 81*(5), 1763–1803.

Taylor-Robinson, D., E. Rougeaux, D. Harrison, M. Whitehead, B. Barr, and A. Pearce (2013, 12). The rise of food poverty in the UK. *British Medical Journal 347*.

Train, K. E. (2003). *Discrete Choice Methods with Simulation.* Cambridge University Press.

Treasury, H. (2010). Innovation Index 2009; Innovation, knowledge spending and productivity growth in the UK.

Triplett, J. (2004). Handbook on hedonic indexes and quality adjustments in price indexes: Special application to information technology products. Technical report, OECD Science, Technology and Industry Working Papers 2004/9.

US Department of Agriculture (1997). Understanding the food choices of low-income families. Summary of Findings. Report by Steven Bradbard, Eileen F. Michaels, Kathryn Fleming and Marci Campbell.

US Department of Agriculture (2000). The effect on dietary quality of participation in the food stamp and WIC programs. http://www.ers.usda.gov/publications/fanrr-food-assistance-nutrition-research-program/fanrr-9.aspx, accessed 27 November 2014. Report by Parke E. Wilde, Paul E. McNamara, and Christine K. Ranney/.

US Department of Agriculture (2007). Development and evaluation of the healthy eating index-2005. Report by Patricia M Guenther, Jill Reedy, Susan Krebs-Smith, Bryce Reeve and Peter Basiotis for the Center for Nutrition Policy and Promotion.

US Department of Agriculture (2010). Food expenditures and diet quality among low-income households and individuals. Report by James Mabli and Laura Castner and James Ohls and Mary Kay Fox and Mary Kay Crepinsek and Elizabeth Condon. http://www.fns.usda.gov/food-expenditures-and-diet-quality-among-low-income-household-and-individuals, accessed 31 October 2014.

US Department of Agriculture (2013). Household Food Security in the United States in 2012. Economic Research Service Report 155, by Coleman-Jensen, Alisha, Mark Nord and Anita Singh.

US Department of Agriculture (2014). Economic research service food expenditures. http://www.ers.usda.gov/data-products/food-expenditures.aspx, accessed 30 October 2014.

Villas-Boas, M. (1993). Predicting Advertising Pulsing Policies in an Oligopoly: A Model and Empirical Test. *Marketing Science 12*, 88–102.

Villas-Boas, S. B. (2007). Vertical relationships between manufacturers and retailers: Inference with limited data. *The Review of Economic Studies 74*(2), 625–652.

Weyl, E. G. and M. Fabinger (2013). Pass-through as an economic tool: principles of incidence under imperfect competition. *Journal of Political Economy 121*(3), 528–583.

WHO (2010). *Set of recommendations on the marketing of foods and non-alcoholic beverages to children.* Geneva: World Health Organization.

Wilson, J. D. (1999). Theories of tax competition. *National tax journal*, 269–304.