cemmap
centre for microdata methods and practice

Estimating households' willingness to pay

Rachel Griffith
Lars Nesheim

The Institute for Fiscal Studies
Department of Economics, UCL

cemmap working paper CWP24/10

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL
An ESRC Research Centre

Institute for Fiscal Studies, 7 Ridgmount Street, London WC1E 7AE **Tel:** 44 (0)20 7291 4800 **Email:** info@cemmap.ac.uk **www.cemmap.ac.uk**

Estimating households' willingness to pay

Rachel Griffith
Institute for Fiscal Studies and UCL

Lars Nesheim
CeMMAP, Institute for Fiscal Studies and UCL

July 22, 2010

Abstract

The recent literature has brought together the characteristics model of utility and classic revealed preference arguments to learn about consumers' willingness to pay. We incorporate market pricing equilibrium conditions into this setting. This allows us to use observed purchase prices and quantities on a large basket of products to learn about individual household's willingness to pay for characteristics, while maintaining a high degree of flexibility and also avoiding the biases that arise from inappropriate aggregation.

We illustrate the approach using scanner data on food purchases to estimate bounds on willingness to pay for the organic characteristic. We combine these estimates with information on households' stated preferences and beliefs to show that on average quality is the most important factor affecting bounds on household willingness to pay for organic, with health concerns coming second, and environmental concerns lagging far behind.

JEL: D12, L11, L81, Q51, C81

Correspondence: rgriffith@ifs.org.uk, l.nesheim@ucl.ac.uk

Acknowledgement: The authors would like to thank James Banks, Richard Blundell, Martin Browning, Ian Crawford, Andrew Leicester, Aviv Nevo, Ariel Pakes and Carol Propper for many helpful comments. Financial support from the ESRC through the ESRC Centre for the Microeconomic Analysis of Public Policy at IFS (CPP) and the ESRC Centre for Microdata Methods and Practice (CeMMAP) is gratefully acknowledged. All errors remain the responsibility of the authors.

1 Introduction

Researchers often want to estimate the value that consumers place on a specific characteristic. For example, to evaluate the consumer benefits of organic farming, researchers need an estimate of willingness to pay for the organic characteristic in organic products. When the number of relevant products is small, or when utility is a separable function of a small number of aggregates, hedonic price or discrete choice demand methods can be used to estimate this willingness to pay. However, these methods cannot be easily adapted to estimate willingness to pay when a consumer chooses to buy hundreds of items from a choice set with tens of thousands of options. When consumers buy a large basket of goods and separability is not invoked classical demand methods are required to model interactions across goods. In an important recent contribution, Blow, Browning and Crawford (2008) improve our ability to estimate consumers' willingness to pay by embedding a characteristics model of utility in a classic revealed preference setting. However, their methods remain intractable when there are more than a few relevant characteristics. What is needed is a method to marry hedonic price methods to revealed preference methods for analysing baskets of goods.

We combine methods from the hedonic pricing literature with revealed preference ideas from Blow, Browning and Crawford (2008) to show how these can be applied to analyse willingness to pay when consumers purchase continuous quantities of a high dimensional basket of goods. Thus, we show how disaggregate analysis of a seemingly impossibly high dimensional dataset can be made tractable. Further, we extend the revealed preference approach by incorporating market pricing conditions. We demonstrate how this approach allows us to learn about consumers' willingness to pay for characteristics using feasible methods, while avoiding unnecessary separability assumptions and aggregation bias. A major benefit of our approach is that we can examine rich data with minimal assumptions. A cost is that we are only able to estimate bounds on willingness to pay.

We apply this approach to an issue of empirical interest; we estimate bounds on willingness to pay for organic foods and show how these are related to demographic characteristics and to consumers' stated preferences and beliefs about health, the environment and product quality. These estimates can inform regulation over the licensing and labelling of organic foods, increase government knowledge about consumer valuations of agricultural and environmental policies, and help give firms a better understanding of the potential profitability of new product lines.

In the consumer demand literature, our work is most directly related to Blow, Browning and Crawford (2008), who develop non-parametric revealed preference methods to estimate willingness to pay for characteristics and apply them to organic milk sales in Denmark. We discuss this paper and its relation to our work in section 2.5. In the hedonic literature, our work is most closely related to papers that focus on discrete choices in imperfectly competitive markets such as Pakes (2003), Bajari and Benkard (2005a,b), and Erickson, Pakes, and Center (2008).[1] Also related is the industrial organisation literature that applies discrete choice demand methods to model demand for single categories of products.[2]

Our work is also related to the price index literature. The most closely related papers are Hausman (2003), Pakes (2003), and Triplett (2005). Triplett (2005) presents a comprehensive discussion of uses of hedonic methods in constructing price indexes. Effectively, what we do is compute consumer specific hedonic price indexes and analyse the implications of these for consumer valuation of organic foods. Although there is no simple relationship between hedonic prices and consumer valuation, since the hedonic price function is determined by the interaction of demand, cost and competitive conditions (Hausman (2003)), we exploit the idea that the hedonic price is a lower bound to compensating variation.

It has long been understood that analogues of classic revealed preference arguments apply to hedonic prices (see for example Scotchmer (1985), Kanemoto (1988), Pollak (1989), and Pakes (2003)). These papers show that hedonic prices can be used to bound willingness to pay

[1] Also related are papers studying hedonic prices in competitive markets in labor economics (Sattinger (1995), Leeth and Ruser (2003)), in environmental economics (Freeman (1995), Smith and Huang (1995), Chay and Greenstone (2005), Sieg et al. (2004)) and urban and public economics (Epple and Sieg (1999)).

[2] See, for example, Berry, Levinsohn and Pakes (1995), Nevo (2001), Smith (2004), and Griffith, Nesheim, O'Connell (2010). Also related is the discrete choice demand literature, e.g. McFadden (1974) and Train (2003).

and willingness to accept. We build on Scotchmer (1985) and Pollak (1989) to develop the argument when consumer choice involves choice of a discrete option along with a continuous intensity of use for a basket of goods. The fact that a consumer paid a premium to purchase a basket of goods implies that the consumer must have been willing to pay at least as much as the premium.[3] Our work is also inspired by the growing interest in partially identified models (e.g. Manski (2003), Chernozhukov, Hong, and Tamer (2007)).

We are in part motivated by the fact that detailed transaction level data on consumer expenditures - which include precisely measured item level prices, quantities and product characteristics - are now widely available for many different goods and in many countries.[4] In our case, our sample is representative of the entire retail grocery market in the UK. The richness of these data are both a blessing and a curse. Discrete choice demand methods, are not tractable for such a large number of product groups. Classical demand methods cannot cope with the large number of zeros in the consumer demand system. Methods such as those introduced by Blow, Browning and Crawford (2008) are not able to deal with the large dimensionality of the characteristic space.

In our application, we use data on each item in the household's food basket, to estimate bounds on willingness to pay for the organic characteristic. We estimate these bounds both for individual product categories and for the entire household food basket. The former bound provides evidence on the importance of organic for individual product lines. The latter provides evidence on the overall consumer valuation properly weighted by expenditure shares.

We show that there is substantial heterogeneity in organic price premia across products

[3] The bound is not structural except under very restrictive assumptions. It may change when market prices change. To estimate structural demand parameters or supply parameters one must use techniques such as those in Epple and Sieg (1999), Ekeland, Heckman, and Nesheim (2004), Smith (2004), Bajari and Benkard (2005a), or Heckman, Matzkin and Nesheim (2010).

[4] For example, data on grocery purchases are available from market research firms, such as AC Nielsen and Kantar (previously TNS), in the US, Canada, the UK and many European countries. Work using these data has looked either at the aggregate basket of groceries (e.g. Smith, 2004) or at single product categories, for example, breakfast cereals (Nevo, 2001), ketchup (Pesendorfer, 2002), yoghurt (Ackerberg, 2001) or carbonated soft drinks (Dube, 2005).

and across households. For products, organic price premia range from -41% to 142% with most of the price premia (80% of them) ranging from 15% to 70%. For households, lower bounds on willingness to pay for the entire basket range from 0 to over £100 per year. 23% have a lower bound of zero and nearly 10% have a lower bound larger than £10 per year.

We also find that consumers vary in the reasons they are willing to pay for organic products, with product quality and health being the most important factors and environmental concerns lagging far behind. Aggregating our results, we estimate that the total lower bound on willingness to pay for health is around £17m, for the environment around £5m, and for quality around £18m. These results have implications for the regulation of organic labelling, and for the way that firms may want to advertise organic products.

The rest of the paper is structured as follows. The next section describes the theoretical background. Section 3 describes the data, our empirical implementation and presents estimates of the lower and upper bounds on households' willingness to pay for organic. Section 3.4 considers the extent to which this willingness to pay reflects concerns about the environment, health or quality and analyses some implications of the results. A final section summarises and makes some concluding remarks.

2 Theoretical background

We employ a rational choice model of a price-taking consumer who maximises utility. At each point in time, the consumer chooses a shopping basket given the set of products and stores in the market and the prices of all products. Prices are set by firms in an unspecified pricing equilibrium. The consumer's choice reveals bounds on their willingness to pay for specific characteristics available in the market.

Our approach to identifying a lower and upper bound on willingness to pay requires the following assumptions, which are common to most of the applied demand literature:

1. Utility depends on observable characteristics that are well measured in our data and possibly on unobserved characteristics that are mean independent of the observed char-

acteristics.

2. Consumers maximise utility, have complete knowledge of the market environment and incur no search frictions.

3. The market sets an equilibrium price for each product and the set of marketed products is rich enough to allow consumers to effectively unbundle product characteristics.

To fix ideas, we follow Blow, Browning and Crawford (2008) and use the example of the organic characteristic, but, the approach could be applied to any other characteristic, for example brand name or whether a product is low-fat. To develop intuition we first describe the single product case. We then describe the market environment, before extending the analysis to the choice of a basket of products, some fraction of which have the organic characteristic, to show how estimates of hedonic price premia from disaggregate purchase data can be correctly aggregated.

2.1 Demand for a single product

In the single product discrete choice model, a consumer maximises utility by choosing one product from a finite number of options, each described by a vector of characteristics and a price. Let $z \in Z \subseteq \mathbf{R}_n$ be the vector of *all* product characteristics that affect consumer choice.[5] Let $z(j)$ be the $j'th$ coordinate of the characteristic vector and let $z(1)$ be the relevant characteristic of interest. In our example, $z(1) = 1$ if a product is organic and $z(1) = 0$ otherwise. The product price is given by $p = h(z)$, where $h(z)$ is an equilibrium hedonic price defined for all $z \in Z$.[6] We discuss the market environment and determination of prices in section 2.2.

[5] To simplify notation in this section, we use z to represent the vector of all product characteristics. Later, in Section 3.2, we change notation slightly and use z to represent the vector of observable product characteristics and ε to represent unobservable characteristics.

[6] Note that the hedonic price is defined for all $z \in Z$ even those that are not sold in equilibrium. See Ekeland (2010) or Chiappori, McCann, and Nesheim (2010) for a discussion of equilibrium pricing of products that are not sold.

Consider a consumer (indexed h for household) with characteristics x_h who buys a single unit of an organic product o with product characteristics z^o and price p^o and elects not to buy a non-organic product n with characteristics z^n and price p^n. The vector x_h includes measured characteristics as well as unmeasured heterogeneity and preference shocks such as the arrival of household visitors, weather, travel cost shocks, or random variation in who within the household does the shopping. Note that our specification encompasses standard discrete choice models such as the mixed logit.

Assume that o and n are identical in all dimensions other than organic. Let the consumer's indirect utility function be denoted $v(x_h, z, p)$, where v is increasing in $z\,(1)$, continuously differentiable in p and strictly decreasing in p. If the consumer chooses the organic product, then standard revealed preference arguments imply

$$v(x_h, z^o, p^o) \geq v(x_h, z^n, p^n), \tag{1}$$

the consumer obtains weakly greater utility from the organic product. By the mean value theorem, there exists some $p^* \in [p^0, p^n]$ such that

$$v(x_h, z^o, p^n) + \frac{\partial v(x_h, z^o, p^*)}{\partial p}(p^o - p^n) \geq v(x_h, z^n, p^n),$$

where $-\frac{\partial v(x_h, z^o, p^*)}{\partial p} > 0$ is the marginal utility of income. After rearranging, we have

$$\frac{v(x_h, z^o, p^n) - v(x_h, z^n, p^n)}{-\frac{\partial v(x_h, z^o, p^*)}{\partial p}} \geq p^o - p^n. \tag{2}$$

The left side of this expression is the willingness to pay for the organic characteristic when we consider a single organic item. The right side is the organic price premium. Revealed preference implies that willingness to pay is at least as big as the price premium. For all consumers that we observe buying organic, the price premium defines a lower bound on their willingness to pay for organic. For all consumers that do not buy organic, the price premium provides an upper bound on their willingness to pay for organic.

To make use of this inequality, we must observe or estimate $p^o - p^n$. We next discuss the market environment that generates prices.

2.2 Market environment

The prices that we observe are equilibrium outcomes in the market, as has been emphasized in much of the recent hedonic and industrial organization literatures (Berry, Levinsohn and Pakes (1995), Ekeland, Heckman and Nesheim (2004), Bajari and Benkard(2004)). For a single food category they are described by the hedonic price function

$$p = h(z).$$

In the markets we consider (UK retail grocery markets), there is effectively no price discrimination conditional on observed product characteristics. On the demand side, consumers shopping in the same market at the same time face the same menu of prices. Similar goods selling for different prices are indeed different goods. For example, two different pack sizes of the same brand of product are distinct because their storeability characteristics vary. The same brand sold in a large, out-of-town store and a convenience store is not the same product because the stores differ in their service and location characteristics.

On the supply side, grocery markets are typified by oligopolistic competition - a small number of firms sell a large number of differentiated products to a large number of consumers. Our approach is consistent with the set of products and the price premia that we observe being determined in an oligopoly marketing and pricing game in which each firm chooses what products to sell and for what price. To study demand we do not need to be specific about the nature of this pricing game, other than to assume that 1) all consumers face the same prices, and 2) the observed market outcomes reflect the technological feasibility of producing and selling various products, the cost structure of firms, the nature of competition in the market, and the distribution of demand across locations and products.

2.3 Demand for a basket

When a consumer buys a basket of goods, some organic and some not, their willingness to pay is defined in terms of the characteristics of all the goods and the total cost of the basket.

Some baskets are more organic than others; they have a larger fraction of items that are organic. We define a non-organic basket to be one in which no products are organic. A fully organic basket is one in which all products are organic.

Let B be the finite set of all products in the market. Partition the set into G categories, with each category labeled by an integer $g \in \{1, ..., G\}$. Let B_g be the set of products in category g and $B = \bigcup_{g \in G} B_g$. Each product $b \in B_g$ has a vector of characteristics $z_b \in Z_g \subseteq \mathbf{R}_{n_g}$ that affect utility. The set Z_g is the set of feasible characteristics for product category g. As in the single product example above, we assume that $z_b(1) = 1$ if and only if b is organic. We define $\overline{z} = \{z_b\}_{b \in B}$ to be the vector of all characteristics of all products.

The price of each product is p_b and the vector of all prices is $\overline{p}$. As discussed in Section 2.2, prices are set in an oligopolistic equilibrium game. For each category g, the price of product b is given by $p_b = h_g(z_b)$ where h_g is the hedonic price function for category g. For each g, the function h_g is defined for all $z \in Z_g$ including those not sold in equilibrium.

It is convenient to work in terms of the consumer expenditure function. Let $\overline{v} = v(x_h, \overline{z}, \overline{p})$ be the maximum utility obtainable given market prices $\overline{p}$ and product characteristics $\overline{z}$. Each consumer chooses a vector of quantities of each product, $\overline{q}_h$ to minimise costs of attaining the fixed utility level $\overline{v}$. The consumer's total expenditure is

$$\begin{aligned} e_h &= c(\overline{p}, \overline{z}, x_h, \overline{v}) \\ &= \min_{\overline{q}} \{\overline{p} \cdot \overline{q} \text{ subject to } u(x_h, \overline{z}, \overline{q}) \geq \overline{v}\}. \end{aligned}$$

In general, the basket purchased will include both organic and non-organic products and the fraction organic will vary across consumers.

For each consumer, we observe the actual basket purchased and the price and characteristics of all items purchased. How do we define willingness to pay for organic? In the discrete choice case, willingness to pay is defined with respect to an alternative that is identical in every dimension except organic. When the consumer purchases a basket, however, there are multiple dimensions of organic, one for each product in the basket. We define willingness to pay to be the difference in expenditure between the amount actually spent and what would have been spent if all organic items were transformed into non-organic while holding **utility** constant. We calculate the lower bound on willingness to pay by comparing the consumer's actual expenditure to what would have been spent if all the products purchased were transformed into non-organic products while holding the **bundle** $\overline{q}_h$ constant.

Formally, let $\overline{z}^n = \{z_b^n\}_{b \in B}$ be a counterfactual vector of characteristics with $z_b^n(1) = 0$ and $z_b^n(j) = z_b^0(j)$ for $j > 1$ and for all $b \in B$. The vector $\overline{z}^n$ is the vector of characteristics in the counterfactual world in which all organic products are transformed into non-organic products. Holding the hedonic price schedule fixed, for all $b \in B$ and for all g counterfactual prices are given by $p_b^n = h_g(z_b^n)$. Let $\overline{p}^n = \{p_{bs}^n\}_{(b,s) \in B \times S}$ is the vector of counterfactual prices. Then, counterfactual expenditure would be

$$e_h^n = c\left(\overline{p}^n, \overline{z}^n, x_h, \overline{v}\right)$$

In this counterfactual, when characteristics are $\overline{z}^n$ and prices are $\overline{p}^n$, the cost minimising basket is $\overline{q}_h^n$.

The total organic expenditure premium or willingness to pay for organic is

$$e_h - e_h^n = WTP_h^n.$$

Note that it is the negative of compensating variation; the amount that exactly compensates a consumer for a change from $(\overline{p}, \overline{z})$ to $(\overline{p}^n, \overline{z}^n)$.

Since the utility function is not known, we cannot calculate this willingness to pay. However, revealed preference gives a lower bound,

$$WTP_h^n = e_h - e_h^n \geq \left(\overline{p} - \overline{p}^n\right) \cdot \overline{q}_h. \tag{3}$$

By choosing to purchase $\overline{q}_h$, the consumer has revealed that they are willing to pay at least $\left(\overline{p} - \overline{p}^n\right) \cdot \overline{q}_h$ to purchase organic. This follows immediately from cost minimisation since

$$\overline{p}^n \cdot \overline{q}_h \geq e_h^n.$$

In words, with counterfactual prices $\overline{p}^n$, the cost of the original bundle is at least as large as the new cost minimising basket. In a similar way, we can use our estimates to calculate bounds on willingness to pay for alternative counterfactual baskets.

We can also compute various upper bounds for willingness to pay by considering counterfactual bundles in which some non-organic products are converted to organic. For example, consider the extreme counterfactual bundle in which all products become organic. Let $\overline{z}^o = \{z_b^o\}_{b \in B}$ be the "all-organic" counterfactual characteristics vector with $z_b^o(1) = 1$ and $z_b^o(j) = z_b(j)$ for $j > 1$ and for all $b \in B$. Let $p_b^o = h_g(z_b^o)$ for all $b \in B$ and for all g. For this counterfactual bundle, we can compute upper bounds on willingness to pay for each consumer using,

$$w_h^o = (\overline{p}^o - \overline{p}) \cdot \overline{q}_h.$$

This characteristics bundle is the maximally organic bundle; all products are transformed into organic products.

In summary, for each consumer we can calculate lower and upper bounds on willingness to pay for organic using

$$w_h^n = (\overline{p} - \overline{p}^n) \cdot \overline{q}_h \leq e_h - e_h^n \tag{4}$$

$$w_h^o = (\overline{p}^o - \overline{p}) \cdot \overline{q}_h \geq e_h^o - e_h. \tag{5}$$

For each consumer that purchases any products with the organic characteristic, equation (4) provides a lower bound on willingness to pay for the bundle of organic items actually purchased. For each consumer that purchases any non-organic items, equation (5) provides an upper bound for willingness to pay for organic for all non-organic items purchased.

For a single household, these lower and upper bounds are only comparable under very restrictive conditions - if all goods are strictly separable in utility and there are no time varying preference shocks. In this special case, each purchase event (one Cox apple on Tuesday, one Tesco 2 litre whole milk on Wednesday, one can of beans on Thursday, etc) provides information on the same willingness to pay for the organic characteristic. It is easy to see that the data are not consistent with this special case. If we attempt to impose these conditions, nearly all households violate the conditions of revealed preference.

More generally, the bounds are not comparable because willingness to pay for each component of the organic basket depends on the vector or characteristics purchased and on time varying preference shocks. The lower bound is a lower bound for willingness to pay for the organic characteristic for the fraction of the basket that is represented by products with the organic characteristic, while the upper bound is an upper bound for willingness to pay for the fraction of the basket that is represented by products that do not have the non-organic characteristic. To estimate the missing bounds - an upper bound for the organic fraction of the basket and a lower bound for the non-organic fraction - we would need to add further assumptions on preference shocks and store switching behaviour or use the panel aspect of the data to identify willingness to pay.[7] What we do in this paper is consider what we can learn by exploiting revealed preference and equilibrium pricing alone.

In comparisons across households, it is important to note that the price vectors are the same across households while the baskets vary. Our measured bound on willingness to pay is the revealed price index for organic characteristics in the consumer food basket. It is a Laspeyres index. It reveals first order bounds on WTP, but in common with most price indices it does not account for potential substitution that might occur if prices did indeed change.

2.4 Household heterogeneity

The expenditure premia in (4) and (5)are consumer specific. While the prices are the same for all consumers, the baskets chosen are not. Consumers make different purchase choices. Hence, the expenditure premia vary across consumers.

In our application we consider how our estimated lower bounds vary with demographic

[7] In ongoing work we are estimating a discrete-continuous demand model that imposes further structure on utility, exploits the panel nature of the dataset and exploits household level price variation across transactions induced by random shocks to the store choice process. Exploiting the repeated observations in the panel data is more complicated than in a simple discrete choice framework, for example, because the dimension of the vector $\overline{q}_{ht}$ is very large.

characteristics and in particular with stated preference survey responses to attitudinal questions that capture some aspects of the main shopper's preferences and beliefs. Our estimated lower bounds are

$$w_h^n = (\overline{p} - \overline{p}^n) \cdot \overline{q}\,(x_h, \overline{\eta}_h)$$

where we have written $\overline{q}\,(x_h, \overline{\eta}_h)$ to emphasise that consumer demand depends on observable consumer characteristics x_h and unobservable consumer characteristics $\overline{\eta}_h$. Without imposing more structure and/or using the panel nature of the data, we cannot estimate the demand functions $\overline{q}$. In particular, we expect that quantities demanded will depend on prices and that prices will depend on unobserved heterogeneity. However, we can estimate a reduced form willingness to pay regression of the form

$$\ln w_h^n = d\,(x_h) + \eta_h, \tag{6}$$

where x_h is a vector of consumer characteristics, as described above, and η_h is a scalar error term. This regression gives us some idea of how important different observed consumer characteristic are in explaining variation in consumer level lower bounds on willingness to pay for organic products.

2.5 Comparison to literature

The hedonic and industrial organisation literatures have largely focused on the analysis of discrete choices over single commodities.[8]. This work does not typically use information that is available on the intensive margin of quantity purchased nor information on simultaneous purchases of multiple discrete commodities with the same characteristic. Methods have been developed to analyse these cases. Scotchmer (1984), Dubin and McFadden (1984), Haneman (1984), Smith (2004), Beckert, Griffith and Nesheim (2009) analyse demand with both discrete and continuous margins and Hendel (1999) and Dube (2005) analyse multiple dis-

[8] Examples include: ready to eat cereals (Nevo, 2001), cars (Berry, Levinsohn and Pakes (1995, 2004), computers (Bajari and Benkard (2005a, b)), butter and margarine (Griffith, Nesheim, O'Connell, 2010), houses Chay and Greenstone (2005) or jobs Leeth and Ruser (2004).

crete choices. However, all of these methods remain intractable in cases where the consumer basket contains a large number (i.e. dozens or hundreds) of products.

In order to study demand for the entire basket it is useful to move back towards classical demand methods that were developed to analyse demand for a basket of goods. However, the number of products that can be feasibly studied is limited here as well. Parametric classical demand methods generally aggregate goods. Similarly, revealed preference methods have typically aggregated goods up to a small number of commodities.[9] Blow, Browning and Crawford (2008), henceforth BBC, extended classical revealed preference methods to study willingness to pay for characteristics of goods. However, the number of characteristics that can feasibly be studied in their set up is small. BBC impose that milk is separable from other goods and study willingness to pay for organic milk. They state *"if we drop separability, then we are left with an impossibly wide problem (hundreds of goods and dozens of characteristics.)"* If we want to look at something like the basket of food purchased in supermarkets the scale of the estimation problem is beyond these methods. Our paper shows that it is possible to analyse the *"impossibly wide problem"* if one is willing to study bounds on willingness to pay. To clarify some of the empirical issues involved, we discuss BBC and its relationship to our work in more detail.

2.5.1 Blow, Browning, and Crawford (2008)

BBC embed a characteristics model of utility in a classic revealed preference setting, thus bringing together the classical demand literature with the literature on hedonic or characteristics models. Their most general theoretical model is indeed very general. It is equivalent to our model except in two dimensions: it does not allow for time varying demand shocks, and it abstracts from the process by which prices are set.

In addition, a number of differences arise in comparing their empirical application with ours. First, BBC study a single product, milk; we study the entire basket of groceries. Second, they impose additional structure on the problem in their empirical application by

[9] For example, see Blundell, Browning and Crawford (2003).

considering only a subset of the relevant product characteristics. They assume that the relevant characteristics for milk are “milkiness", three levels of fat, and organic; thus assuming that there are six types of milk. In Table 3 we show that several other characteristics are important for describing price variation, including pack size, fascia (store format), packaging, and filter method. Ignoring these other characteristics results in biased estimates of the willingness to pay for the organic characteristic. Further, when combined with aggregation, it results in spurious variation in prices across consumers and across time.

BBC start with data that is equivalent to ours (transaction level data) but aggregate milk purchases within households across time, across pack sizes and across stores to compute unit prices or “unit values" of six types of milk.[10] These “unit values” vary within category reflecting differences in the products purchased, as well as differences in the prices faced by households (see, inter alia, Lahatte et al (1998), Deaton (1987, 2006)). This aggregation causes several problems. First, the aggregation can lead to a failure to reject rationality when it should be rejected. Second, the variation in prices across households and across time is spurious. In the cross section, all households shopping in the same market face one market price. If all characteristics are included, there is no variation in prices. Any dispersion in prices across households is spurious (due to different households buying different sizes or shopping at different stores). Similarly, over time, at least some of the time series variation in prices that BBC use results from households switching stores or pack sizes, due to demand shocks. In our data, most of the time series variation in prices is eliminated after accounting for variations in choices of store and pack size.

Finally, because they are interested in recovering time invariant preferences of consumers, BBC assume that there are no demand shocks. Classical revealed preference methods require this assumption to non-parametrically recover preferences from time series of consumer demands. However, in common with the discrete choice demand literature we believe it is quite natural to assume that day-to-day purchases are affected by demand shocks. At the disag-

[10] BBC use data for Denmark over the period 1999-2000. We use data for the UK over 2004.

gregate level, it is not possible to explain consumer purchases without allowing for demand shocks.

We illustrate these points with an example. Consider a consumer for whom we observe the following four items purchased:

Date	Pack size	Store	Product	price
1 March	pint	Sainsbury	Organic skimmed	1.40
6 March	half gallon	Sainsbury	Nonorganic skimmed	4.40
9 March	pint	Tesco Metro	Nonorganic skimmed	1.60
14 March	pint	Sainsbury	Organic skimmed	1.50

BBC assume that there are only six milk products defined by fat content (low, medium and high fat) and whether or not organic. They then aggregate data to the household-month-product level, calculating an average unit price for each household and for each product (across other characteristics). From the information in the table above, BBC would conclude that:

- for this consumer, the unit price of organic skimmed milk is $(1.40+1.50)/2 = 1.45$ per pint and the unit price of non-organic skimmed milk is $(4.40+1.60)/5 = 1.20$ per pint.
- this consumer's willingness to pay for organic is 0.25 per unit of organic.

If we had only observed this consumer making the 1 March and 9 March purchases, BBC would conclude that this consumer failed the General Axiom of Revealed Preferences (GARP). Because they purchased non-organic milk at a higher price than organic, the information on this consumer would have been discarded.

How does this compare to what we do? Instead of computing 6 prices at each point in time, we compute 535 prices.[11] In the cross section, we observe little variation in prices that is not explained by product characteristics. In the panel, we observe some but not much

[11] There is one price for each combination of characteristics that we observe in the market. The relevant characteristics include store (7 levels), organic (2 levels), pack size (4 levels), own branded (3 levels), fat content (3 levels), packaging (4 levels), variety (6 levels) and filtering method (5 levels). Not all possible combinations are observed. We also control for region within the UK and whether the product is on ticket price reduction, multi-pack offer, or extra-free promotion.

variation. In contrast, BBC obtain variation in their measured organic price premium both across individual consumers and across time. Where does this variation come from? It arises (in part at least) because the products purchased vary across consumer and across time; BBC aggregate over products and time. In the example above, the price of milk purchased on 1 March differs from that purchased on 6 March not only because it is organic, but also because of the pack size. It is well known that larger pack sizes are priced more cheaply per volume.[12] Omitting package size, which in this example is correlated with organic, leads to an over-estimate of the organic price premium.

A final difference that is worth pointing out is that BBC do not use information on households that either never or always buy organic products. In our data 18.8% of households never buy any products with the organic characteristic (no households buy only organic products.). If we look just at milk, 85% of households do not buy any organic milk while 0.25% buy only organic milk. In both cases, these are important fractions of the population. For those who don't buy, we provide upper bounds to willingness to pay. For those who always buy, we provide lower bounds.

3 Empirical application

To illustrate the approach described above we consider households' willingness to pay for the organic characteristic. This is an interesting application in it's own right and allows a direct comparison to Blow, Browning and Crawford (2008). We start by describing our data. Then we specify the hedonic function and show an example of how we estimate the organic price premia for one food category. Finally, we show summary results from similar hedonic regressions for 75 grocery food categories, and estimate bounds on willingness to pay for organic derived from an analysis of the basket of grocery products.

[12] Feenstra and Shapiro (2003), Triplett, (2003), Hendel and Nevo (2006), and Griffith, Leibtag, Leicester and Nevo (2009).

3.1 Data

Our data come from the TNS/Kantar Worldpanel for calendar year 2004. The data are representative of the entire UK retail grocery market. Households record purchases of all items that are brought into the home using hand-held scanners. Prices are recorded from till receipts collected from the households. We use information on prices, quantities and characteristics of food items purchased for home consumption by 16,881 households. The sample contains data on more than 11 million purchases. The characteristics data includes information on a large number of product characteristics judged to be important by market researchers, as well as store identity. Demographic information and information on a range of household attitudes (including attitudes towards health, quality, the environment and organic) is collected annually by a telephone survey.[13] We have sampling weights that allow us to gross-up to population figures.

Individual food products (defined by a unique bar code) are grouped into categories such as "fresh lamb", "tea", "olives", etc. We use data on 75 categories where organic is a relevant characteristic. Total expenditure on these products in our sample of households is £12.8m (grossed up using sampling weights it is £19.7bn). The 75 food categories are shown in Table 1 along with number of purchases, grossed up and observed expenditure and the share of organic. On average 2.1% of expenditure is on products that have the organic characteristic. This varies substantially across categories ranging from 0.4% of "Fresh Bacon Rashers" to 28.6% of "Chilled Meat and Vegetable Extract."

Table 2 shows organic purchases at the household level. Just under 20% of households never buy any organic products and over a quarter buy only a very small amount (less than half of one percent of their total expenditure). However, 37% of households spend more than 1% of their budget on organic products, and there are a small number of households (7%) that spend over 5% of their budget on organic products. These numbers illustrate the tremendous heterogeneity in demand for organic products, and that organic is an important

[13] See Leicester and Oldfield (2009) and Griffith and O'Connell (2009) for further details on the data.

expenditure category for a significant part of the population.

3.2 Hedonic model

For each product category g, we estimate a parsimonious log-linear model with interactions between the organic characteristic and dummy variables representing the dominant supermarket chains. The model includes a large vector of characteristics z as well as dummy variables indicating month, special offer, region and store type (store types include Asda, Marks & Spencer, Safeway, Sainsbury, Tesco, Waitrose and other). These are the main variables driving variation in prices. Since nearly all elements of z are discrete, and since we have included a large set of characteristics, these specifications are very flexible and capture most of the variation in prices.[14]

Let (b, r, s, t) index items, regions, store types and time. For each product category, we estimate a hedonic regression of the form

$$\ln(p_{brst}) = \alpha_1\delta_t + \alpha_2\kappa_{bt} + \alpha_3\phi_r + \beta z_{bs} + \varepsilon_{brst} \tag{7}$$

where δ_t is the vector of eleven monthly time dummies, κ_{bt} is a vector of indicators for special offers (ticket price reduction, multi-pack purchase and extra free), and ϕ_r is a vector of regional dummies (North East, North West, Yorkshire and Humber, East Midlands, West Midlands, East of England, London, South East, South West, Wales, Scotland). The residual ε_{brst} captures unobserved product characteristics that are mean independent of the observed characteristic.

It has long been recognised that interpretation of the error term in hedonic regressions plays an important role in hedonic analysis.[15] In the literature, three main sources of error have been considered (See Triplett (2005) for a detailed discussion.). First, the error term

[14] For each regression, we report the adjusted R^2 in Table 4. In principle, we could estimate a non-parametric hedonic function h_g for each product category g. However, the characteristics vector z is high dimensional. Experimentation with more flexible specifications (for example, Box-Cox and specification with further interactions between discrete characteristics) did not result in qualitatively different results. Details are available from the authors on request.

[15] See, inter alia, Griliches (1961), Epple (1987), Pakes (2003), Triplett (2005), Bajari and Benkard (2005a), and Erickson and Pakes (2007).

could come purely from error in measurement of prices. In this case, if the measurement error is mean independent of the observed characteristics, the estimated hedonic price function is consistent and counterfactual hedonic prices can be predicted using the hedonic price function and ignoring the error term.

Second, it could reflect unmeasured product characteristics that are observed by buyers and sellers and hence that affect market prices. In this case, if the unmeasured product characteristics are mean independent of the observed characteristics we can still estimate the hedonic price function consistently, but we must use care in predicting counterfactual prices. Counterfactual prices rely on an assumption about what value of unmeasured characteristics is assumed for the counterfactual goods. Alternative counterfactuals can be generated under different assumptions about the level of unobserved characteristics that is forecast for the counterfactual good. Transforming a good from organic to non-organic holding everything else constant requires holding the unobservable constant as well.

Third, the error term in the regression could reflect pricing errors. In this case, the analysis is similar to that in the unobserved characteristics case. Alternative counterfactuals can be generated under different assumptions about the level of the "pricing error" for the counterfactual product. In our analysis, we include the error term in our counterfactual predictions because we believe that in our data pure measurement error in prices is relatively minimal while the other two considerations may be more important.

For any particular food category, the predicted organic price premium of a specific product is given by

$$\begin{aligned} \Delta p_{brst}^{n} &= \exp\left(\widehat{\alpha}_1\delta_t + \widehat{\alpha}_2\kappa_{bt} + \widehat{\alpha}_3\phi_r + \widehat{\beta}z_{bs} + \widehat{\varepsilon}_{brst}\right) \\ &\quad - \exp\left(\widehat{\alpha}_1\delta_t + \widehat{\alpha}_2\kappa_{bt} + \widehat{\alpha}_3\phi_r + \widehat{\beta}z_{bs}^{n} + \widehat{\varepsilon}_{brst}\right). \end{aligned} \tag{8}$$

Differences in the observed price for the organic characteristic across products, locations and time reflect the technological feasibility of producing and selling organic products, the cost structure of firms, the nature of competition in the market, and the distribution of demand across locations and products. For example, if the cost differential between organic and non-organic beef is larger than that between organic and non-organic chicken, then, all else equal, the organic price premium on beef will be higher than on chicken. Alternatively, if demand for organic beef is more price elastic than demand for organic chicken, or entry into organic beef production is more elastic to profits, then, all else equal, the organic price premium on beef will be lower than on chicken. In general, for each product category, each of these factors plays a role in determining the equilibrium hedonic price of the organic characteristic.

One potential data issue that has received considerable attention in the revealed preference literature is that p_{brst} is only recorded if item b is purchased in region r at store s at time t. Otherwise, it is not observed. Let $d_{brst} = 1$ indicate that we observe at least one occurrence of the price. We assume that

$$E\left(\varepsilon_{brst}\,|z_{bs}, d_{brst} = 1\right) = E\left(\varepsilon_{brst}\,|z_{bs}, d_{brst} = 0\right) = 0. \tag{9}$$

That is, we assume that the mean log price of unobserved characteristics is no different amongst items not purchased. In our application, this is a weak assumption for several reasons. First, the weighted Kantar/TNS sample is nationally representative of both all households and all expenditure items. By construction, the sample is meant to have the desired property. Second, the sample is large and high frequency. Regularly purchased items appear in the sample with high probability. Infrequently purchased items are items for which a very small fraction of the market has willingness to pay larger than the price. But, this does not necessarily imply that these products have log prices that are higher than average. It could be that log prices are lower than average but that willingness to pay is even lower. In our application, there is no reason to expect that average log prices amongst these items is systematically different from log prices of sampled items. In other applications, this might not hold. For example, Erickson, Pakes, and Center (2008) find that a similar assumption does not hold in monthly data for the television market. In contrast to their study, our data sample has a much higher frequency (daily), focuses on a very different market (groceries), and is a much larger sample of individual transactions.

3.2.1 Results for a single food category: milk

To clarify our empirical approach, and to aid comparison to Blow, Browning and Crawford (2008), we present results for milk in Table 3. In the first column we replicate a specification that is close to Blow, Browning and Crawford (2008). We include in the z vector only the organic and fat content characteristics (interacted), along with common month and region effects. The adjusted R^2 on this regression is small at 0.065. The interactions between organic and fat content are not significant (either individually or jointly) - firms in the UK do not charge differential premia on organic depending on the fat content. Therefore in column 2 we drop these interactions, which changes very little else. In column 3 we add in the full set of characteristics including package size and type, variety of milk, store fascia in which purchased and whether on special offer. Many of these are statistically significant, and the estimated organic premium declines significantly. The adjusted R^2 increases to 0.726 - these additional characteristics explain a substantial proportion of the variation in prices. In the final column we include interactions between the organic characteristic and the store fascia in which the milk was purchased. Across all stores, the average price premium for organic milk is 15% and ranges from 0% at Asda to 13% at Tesco to 30% at Waitrose. Ignoring multiple purchases and quantity choices for the moment, since the market share of organic milk is 2.2%, we can say that roughly 2.2% of households have willingness to pay for organic milk of at least 15% while 97.8% of households are willing to pay no more than 15%. However, those who buy organic milk at Waitrose, reveal a lower bound on willingness to pay of 30%.

3.2.2 Results for all food categories

We repeat this analysis for each of the 75 food categories in the data by running 75 separate regressions of the form of (7). Each regression includes a set of characteristics that is common to all categories (whether the product is organic, whether it is an own-brand product, the store it was sold in (Asda, M&S, Safeway, Sainsbury's, Tesco, Waitrose and other), package size, month and region effects, and whether the product was sold as part of a multi-purchase deal, an extra-free offer or on ticket price reduction (sale)) as well as a set of category specific characteristics. The category specific characteristics vary in number and type. For example, there are over 200 flavours of soup and over 250 flavours of yoghurt. Eggs, on the other hand, have relatively few characteristics - whether they are barn reared or free range, eggs size and whether they are branded.

The organic coefficients and their standard errors along with the adjusted R^2's from the regressions are shown in Table 4. Each row in the table shows results from a separate regression. There is a great deal of variation in the estimated organic coefficients both across product categories and across stores. A histogram of all estimated organic price premia is plotted in Figure 1. Of the 595 potential organic-fascia coefficients we are able to identify 518 (some stores never sell an organic version of some products), 462 are positive and 338 are significantly so (at the 5% level). The unweighted mean of the price premia is 0.40 (suggesting the organic characteristic increases prices by 40%) and the median is 0.38. For each of the major supermarket chains, Figure 2 shows the within-store distribution of price premia across product categories. Asda and Safeway have the smallest mean and median price premia as well as the most categories (8) with price premia less than or equal to zero. Even for these stores, most of the premia are positive and range from zero to 125%. The other stores have higher average and median organic price premia, fewer categories (4 or 5) with non-positive coefficients. In all cases, the range of positive price premia is from zero to nearly 125%. Marks & Spencer has the highest mean and median markup, followed closely by Sainsbury's, Waitrose and Tesco.

The adjusted R^2 are high (with a few exceptions) suggesting that we have captured most of the product characteristics that affect pricing. However, unobserved factors still play a role. Bajari and Benkard (2005b) note that a hedonic price index can ignore the pricing of unobserved characteristics if the relationship between observed and unobserved is stable. We argue that this is the case in our application. In UK retail grocery markets technical change is relatively slow; new and exiting product have small aggregate market shares. We have detailed information on all product characteristics judged to be important by market research firms, including characteristics that vary over time (such as being on special offer) and space (such as being sold in a different store). As indicated by the adjusted R^2, measured characteristics explain most of the variation of prices in our data. Because of the stability of the market, it is quite plausible that the relationship of any unmeasured characteristics to measured is stable.

The organic price premia, combined with the decision to buy or not, gives us bounds on willingness to pay for individual organic items. These item specific bounds can be combined with data on quantities purchased to estimate household specific bounds on willingness to pay for baskets of organic products.

3.3 Bounds on individual households' willingness to pay

The lower bound on an individual household's willingness to pay for organic foods is given by equation (4). For each household, we measure q_{hbrst} as the total quantity of item b purchased at store s in region r by household h in month t, and p_{brst} as the price of item b in store s in region r at time t. The dimension of the vectors $\overline{p}$ and $\overline{q}_h$ are each over 4 million (47,854 barcodes by 12 months by 7 stores). Each element of $(\overline{p} - \overline{p}^n)$ is computed using our estimated hedonic coefficients and equation (8).

Tables 5 shows the distribution across households of the bound in (4) measured in 2004 pounds sterling.[16] Over 20% of households either bought no organic products, or bought only a small amount whose price premium was below zero, revealing that their willingness to pay for organic may very well have been zero or negative. The remaining 80% of households were willing to pay at least some positive amount for products with the organic characteristic. Around one quarter have a lower bound less than a pound a year, while over half were

[16] For households that are not observed throughout the entire year, we gross the lower bound up to the level of a full year using weights that represent the average share of expenditure in each month. For example, if we observe the household from January to October, then we divide the estimated lower bound by the average share of total annual expenditure that is accounted for by those months.

willing to pay more than £1 a year. Around 10% were willing to pay £10 or more and 123 households were willing to pay more than £100 a year. Table 6 expresses these numbers as a share of households' annual expenditure on organic foods. These are household level weighted average organic price premia. We see that almost half of households are willing to pay 20% or more for the organic characteristic on these products.

We also compute estimates of the upper bound on household's willingness to pay for organic based on equation (5). Table 7 shows the distribution of these as a share of total expenditure on non-organic items. Most households would not pay more than 40%, and a substantial proportion would not pay more than 30%, on these items.

As we noted at the end of section 2.3, the willingness to pay varies across households because different households buy different baskets. Different baskets may have different organic contents (e.g. households may buy different organic items or the fraction of items that are organic may differ). The measure we obtain is the total price premium paid for all organic characteristics purchased in the entire food basket.[17]

It is worth emphasising that without making further assumptions our lower bound and our upper bound are not strictly comparable, because they are literally bounds on willingness to pay for organic apples and oranges. However, they are comparable under two conditions. First, suppose indirect utility is separable and takes the form

$$v\left(x_h, g_1\left(z_1, p_1\right), ..., g_B\left(z_B, p_B\right)\right).$$

where (z_b, p_b) is the price and characteristics vector for item b. Then, for each good, the trade-off between characteristics and price is independent of all other goods. Second, suppose the functions g_b are identical for all b. Then the trade-off is the same for all b. Under these two conditions, household x_h will have a unique willingness to pay for organic - a willingness to pay that is the same for all goods. Under these conditions, every organic purchase decision is independent and identical; there is a single threshold. Unfortunately, these assumptions

[17] Comparing across baskets is analogous to comparing across jobs when calculating the wage premium for all job related risks in the value of statistical life literature (e.g. Viscusi and Aldy (2003)).

are unlikely to hold. Many of the food products we study are either close substitutes or complements. Additionally, the fact that the price premia vary across products rejects the second condition.

3.4 Analysis

In this section we discuss two examples that illustrate how these bounds are informative. We first study how the lower bound on willingness to pay varies with household characteristics, beliefs and attitudes, and discuss how this illuminates the reasons that the organic characteristic is valued. Secondly, we evaluate what our bounds can tell us about the potential revenue that a store could earn from introducing a new organic product line.

3.4.1 Reasons for heterogeneity in willingness to pay

Why are households willing to pay for organic food? We combine the estimates presented above with demographic information and survey response data on attitudes towards health, the environment and product quality as described in section 2.4 to shed light on this question. We consider how our estimated lower bound on willingness to pay for organic foods varies with self-reported preferences and beliefs and a number of demographic characteristics. We exploit qualitative survey data that are collected by TNS and consider three main factors that have been highlighted in the literature as being reasons why people value organic, and on which we have data - benefits to the environment, health benefits, and better quality food.

In the survey households are asked to indicate the degree to which they agree with each of the following statements:

1. Organic products are healthier

2. I try to buy a healthy range of foods these days

3. Organic foods are friendlier to the environment

4. I try to buy environmentally friendly products
5. Organic foods are better quality
6. I don't mind paying for quality

For each statement, respondents are asked to choose one response from the list

$$\{\text{Agree strongly, Agree, Neither agree nor disagree, Disagree, Strongly disagree}\}.$$

We treat these responses as qualitative measures of household preferences and beliefs and investigate the statistical relationships between the responses and the lower bounds to household willingness to pay. For each of the three factors (environment, health, quality), we have one response that provides a qualitative measure of beliefs and a second that provides a qualitative measure of preferences.

We first report cross-tabulations of responses to these survey questions, household organic expenditure shares and the lower bounds on willingness to pay for organic. In order to reduce the dimensionality of our tables, we report results that pool the five possible survey responses into two groups, (Agree strongly, Agree) and (Neither agree nor disagree, Disagree, Strongly disagree). Tables 8-10 show these cross-tabulations.

Table 8 shows that 2178 households say both that they try to buy healthy foods and that they think organic products are healthier. For these households organic products make up 4.7% of total expenditure, and these households have a mean lower bound on their willingness to pay that represents 2.2% of their total food expenditure. Agreement with both statements is correlated with high expenditure shares and with high lower bounds on willingness to pay. In contrast, 6649 households do not particularly try to buy healthy foods and do not think organic foods are healthier. Disagreement with both statements is negatively correlated with organic expenditure shares and with the lower bound; organic products make up 1.3% of total expenditure in this group and their estimated lower bound on willingness to pay for organic is 0.3% of total expenditure on food. Tables 9 and 10 display similar figures for the questions related to the environment and quality. A total of 1941 households both feel that organic products are good for the environment and try to buy environmentally friendly products. These households spend 3.7% of their budget on organic foods and are willing to pay at least 0.9% of total expenditure. A smaller number of households, 1334, believe organic products are higher quality and do not mind paying for quality; these households spend a larger share (5.9%) on organic products and have a higher estimated lower bound on willingness to pay of 1.4% of total food expenditure. These tables show that preferences and beliefs related to health, the environment and quality are correlated with organic expenditure shares and with bounds on willingness to pay.

Next we seek to disentangle the relative importance of these beliefs and preferences through a multivariate analysis. We regress the households' lower bounds on willingness to pay on the attitudinal responses discussed above and a range of other household characteristics including: family structure, total annual expenditure on food and non-food items (as a proxy for income), the household's social class, and whether anyone in the household is a vegetarian. The means of these variables are shown in Table 12 (most are discrete variables).

Table 11 reports results from this analysis. The column (1) results are from a model that includes only four dummies, one for agreement with both health statements, one for agreement with both environment statements, one for agreement with both quality statements, and one indicating that the household's responses to these questions were missing. (i.e. the dummies indicate whether the household is in the upper left-hand quadrant of Tables 8, 9 or 10 respectively). Households that care a lot about organic and health have on average a lower bound on willingness to pay that is £5.78 higher than households that do not (i.e. are in any of the other three quadrants of Table 8). Households that care a lot about organic and the environment have a lower bound that is £1.73 higher and those that care about quality have a lower bound that is £8.71 higher. Responses to each of the attitude questions are positively correlated with both organic market shares and lower bounds on willingness to pay for organic. The average lower bound is highest amongst the group for whom quality is important, next highest amongst the “health is important" group, and next highest amongst the "environment is important" group.

Column (2) reports results from a model that adds indicators of households that are in the upper right and lower left quadrants of Tables 8-10. These additional indicators have little effect. Column (3) adds demographic controls. The parameter estimates don't change much and the R^2 increases. Columns (4)-(6) display results from separate regressions for each of the main social classes. There is considerable variation in parameter estimates across the groups. Professional households who care a lot about health have a lower bound on willingness to pay that is nearly twice as high as skilled and unskilled households who care a lot about health (£7.68 versus £4.64 and £3.91). Professional and skilled households who value quality have a lower bound that is nearly twice as high as unskilled households (£8.24 and 9.13 versus £4.96). In contrast, valuation of the environment is similar across the social classes. The estimates in Table 11 allow us to calculate the contribution to willingness to pay lower bounds of each of the three concerns - health, environment and quality. In all cases we see that valuing quality is the characteristic that is associated with households that have the highest lower bound on their willingness to pay, followed by health with environment contributing the least.

If we want to know the aggregate lower bound on the valuation of these three concerns then we need to consider not only the mean lower bound on the valuation for those who have the preferences and attitudes that the particular issue is important, but also the number of households that fall into that group. Combining those two pieces of information we get estimates that suggest that the total lower bound on willingness to pay for health is around £16.9m, for the environment around £5.4m, and for quality around £17.8m. These results are interesting and may be surprising to some people. Quality and health seem to be much more important factors in determining the amount (or at least the lower bound) households are willing to pay for organic products. This has implications for the regulation of organic labelling, and for the way that firms may want to advertise organic products.

3.4.2 Introduction of a new organic product line

The upper bound on willingness to pay is also informative. Suppose a firm were considering whether to convert some set of products to organic. What would be the potential impact? Our results allow us to estimate an upper bound on the potential revenue implications of such a move without imposing strong assumptions. This bound is the firm specific component of our previously computed upper bounds. To illustrate, we compute upper bounds on revenue gains from converting two specific food categories to organic - eggs and vegetables. In a similar way, we could calculate results for any other food category.

In considering such a strategy, three factors drive differences in projected revenue across stores - the baseline expenditure on the category in each store, the current proportion that is organic, and the price premium on organic. We can easily calculate the first two from our data (and they are figures that a store would readily know). To get the third we need to have the hedonic regressions.

Table 13 displays the results. For each store and each product category, columns (2) to (4) show the estimated price premia, expenditure, and the share organic. Column (5) shows our estimate of expenditure if products in the category were converted to organic (assuming no substitution by consumers and no price changes by the firms). Column (6) presents the estimated upper bound on the % increase in revenue from switching all the products in the store to organic.

We find significant variation in the price premia charged by different stores. For eggs, the premia range from 26% to over 49%. Similarly, the organic share of the eggs category varies substantially from 4% in Asda to nearly 20% in M&S. There is also a large variance in the total expenditure on eggs. Putting these together, the upper bounds on the revenue increase from converting all eggs to organic range from about 28% for Asda to a maximum of just over 54% for Sainsbury.

Looking at vegetables (salad and other) loose we see an even wider variance in the price premia on organic across fascia - ranging from 15% to 57%. The organic share ranges from

1% to over 15%. The projected maximum revenue gain from converting all vegetables (salad and other) loose to organic ranges from 15% to over 70% increase.

When considering product line changes, a supermarket could compare these maximum revenue projections to expected costs and begin to make first-order judgements as to which changes might be profitable. These could be used to rule out unprofitable changes and allow the supermarkets to focus more detailed analysis on categories that are potentially profitable. While these estimates do not provide point estimates on revenues or profits, they require very few assumptions about household preferences or behaviour and so are quite robust. They are upper bounds. In particular they ignore substitution effects and competitor responses. Further work would then be needed to estimate more precise consumer substitution responses and to gauge rivals responses.

4 Summary and conclusions

Rich data on spending behaviour are now widely available in a number of countries. These data offer great potential to learn about willingness to pay for many different characteristics. However, their use has been in part thwarted by the sheer scale of the data. Existing revealed preference approaches to estimating willingness to pay can not deal with the large dimensionality of these data.

Methods such as Blow, Browning, and Crawford (2008) illustrate how assumptions about separability and no time varying demand shocks, combined with panel data, can be used to obtain point estimates of willingness to pay, at least for a fraction of the population. We extend the ideas developed in Blow, Browning, and Crawford (2008) by incorporating market pricing equilibrium conditions, which help to reduce the dimensionality of the problem, but allow us to retain much of the flexibility of their approach. We use standard assumptions about market pricing equilibrium and consumer revealed preference behaviour to compute consumer specific bounds on willingness to pay. We show how to aggregate estimates of willingness to pay for individual products across a basket of products in a manner that is consistent with consumer theory. These bounds are Laspeyres style price indexes for differentiated products. In order to recover point estimates of the willingness to pay we would need to make further assumptions about the structure of consumers' preferences. While this is certainly feasible for individual product categories, further work needs to be done to develop a tractable method to analyse the entire food basket.

We illustrate the application of these methods using rich data on households' purchases of food to estimate lower and upper bounds on willingness to pay for the organic characteristic in food. Our results suggest that there is a large amount of heterogeneity in willingness to pay for organic products. We relate these revealed bounds on willing to pay for organic to households' stated preferences over organic products to learn about why households value the organic characteristics. Somewhat surprisingly, quality is the most important reason, health concerns coming second, and environmental concerns lagging far behind. We have also shown how these methods can be applied to calculate an upper bound on the revenue impact to a supermarket of introducing a new product range. These are both applications that have direct practical relevance.

In future work it would be interesting to investigate the panel dimension of our data in order to obtain more precise estimates of structural demand parameters. This will require further assumptions about store choice and of the dynamics of consumer preference shocks.

References

Ackerberg, Daniel (2001), “Empirically Distinguishing Informative and Prestige Effects of Advertising”, *RAND Journal of Economics*, Vol. 32, No. 2.

Bajari, Patrick, and Benkard, C. Lanier (2005a), “Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach", *Journal of Political Economy*, 113(6): 1239-1276.

Bajari, Patrick, and Benkard, C. Lanier (2005b), “Hedonic Price Indexes with Unobserved Product Characteristics, and Application to Personal Computers", *Journal of Business & Economic Statistics*, 23(1): 61–76.

Beckert, Walter, Rachel Griffith and Lars Nesheim (2009), "A Micro-econometric Approach to Geographic Market Definition in Local Retail Markets: Demand Side Considerations", IFS mimeo.

Berry, Steven, James Levinsohn and Ariel Pakes (1995), "Automobile Prices in Market Equilibrium", *Econometrica*, 63(4): 841-890.

Berry, Steven, James Levinsohn and Ariel Pakes (2004), "Differentiated Products Demand Systems from Combination of Micro and Macro Data: The New Car Market", J*ournal of Political Economy*, 112(1): 68-105.

Blow, Laura, Martin Browning and Ian Crawford (2008), “Revealed Preference Analysis of Characteristics Models”, *Review of Economic Studies,* 75(2): 371 - 389.

Blundell, Richard, Martin Browning and Ian Crawford (2003), "Nonparametric Engel curves and revealed preference", *Econometrica*, 71(1): 205-240.

Chay, K.Y., and Greenstone, M. (2005), “Does Air Quality Matter? Evidence from the Housing Market", *Journal of Political Economy*, 113(2): 376–424.

Chiappori, Pierre-Andre, McCann, Robert, and Nesheim, Lars (2010), “Hedonic Price Equilibria, Stable Matching and Optimal Transport: Equivalence, Topology and Uniqueness," *Economic Theory*, 42(2): 317-354.

Chernozhukov, V., Hong, H., and Tamer, E. (2007), "Estimation and Confidence Regions for Parameter Sets in Econometric Models", *Econometrica*, 75(5): 1243–1284.

Deaton, Angus (1987), "Estimation of own and cross-price elasticities from household survey data", *Journal of Econometrics*, 36: 7-13.

Deaton, Angus (2006), "Prices, Unit Values and Local Measurement Units in Rural Surveys: an Econometric Approach with an Application to Poverty Measurement in Ethiopia", *Journal of African Economies*, 15(2): 181-211.

Dube, Jean-Pierre (2005), "Product Differentiation and Mergers in the Carbonated Soft Drink Industry," *Journal of Economics and Management Strategy*, 14(4): 879-904.

Dubin, J. amd D.L. McFadden (1984), "An Econometric Analysis of Residential Electrical Appliance Holdings and Usage", *Econometrica*, 52(2): 345-362.

Ekeland, Ivar (2010), "Existence, Uniqueness and Efficiency of Equilibrium in Hedonic Markets with Multidimensional Types", *Economic Theory*,
42(2): 275–315.

Ekeland, Ivar, Heckman, James and Nesheim, Lars (2004), "Identification and Estimation of Hedonic Models", *Journal of Political Economy*, 112: 60-109.

Epple, Dennis (1987), "Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products", *Journal of Political Economy*, 95(1): 59-80.

Epple, Dennis, and Sieg, Holger (1999), "Estimating Equilibrium Models of Local Jurisdictions", *Journal of Political Economy*, 107(4): 645–681.

Erickson, T., and Pakes, A., and Center, L. (2008) "An Experimental Component Index for the CPI: From Annual Computer Data to Monthly Data on Other Goods", NBER Working

Paper.

Feenstra, R.C. and M.D. Shapiro (2003), “High-Frequency Substitution and the Measurement of Price Indexes”, in Scanner Data and Price Indexes.

Freeman, A.M. (1995), “Hedonic Pricing Methods," in The Handbook of Environmental Economics, ed. Daniel Bromley, Blackwell Publishers Ltd. Oxford, pages 672-686.

Griffith, R., Ephraim Leibtag, Andrew Leicester and Aviv Nevo (2009) "Consumer shopping behavior: how much do consumers save?", *Journal of Economic Perspectives*, Vol. 23, No. 2, pp. 99-120

Griffith, Rachel, Lars Nesheim and Martin O'Connell (2010) "Comparing pass-through of excise and ad valorem taxes in differentiated product oligopoly with multi-product firms: empirical estimates using micro data" IFS mimeo

Griffith, R. and M. O'Connell (2009) "The Use of Scanner Data for Research into Nutrition" *Fiscal Studies*, vol. 30, no. 3/4, pp. 339–365

Griliches, Zvi (1961), “Hedonic price indexes for automobiles: an econometric analysis of quality change," in *The Price Statistics of the Federal Government, NBER Staff Report No. 3,* General Series No. 73, New York: 173-196.

Haneman, W.M. (1984) ‘Discrete/continuous models of consumer demand," *Econometrica*, 52(3): 541-561.

Hausman, Jerry (2003), “Sources of Bias and Solutions to Bias in the Consumer Price Index," *Journal of Economic Perspectives*, 17(1): 23-44.

Heckman, James, Rosa Matzkin and Lars Nesheim (2010), “Nonparametric Identification and Estimation of Nonadditive Hedonic Models, *Econometrica*, forthcoming.

Hendel, I. (1999) "Estimating Multiple-Discrete Choice Models: An Application to Computerization Returns" Review of Economic Studies, 66: 2 (April), 423-446

Hendel, I. and A. Nevo (2006) “Sales and Consumer Inventory”, The RAND Journal of Economics, 37(3), 543-561

Kanemoto, Y. (1988), “Hedonic prices and the benefits of public projects," Econometrica, 56: 981-989.

Lahatte, Agenor, Ruth Miquel, François Laisney, and Ian Preston (1998) “Demand systems with unit values: A comparison of two specifications" Economics Letters Volume 58, Issue 3, 1 March 1998, Pages 281-290

Leeth, J.D., Ruser, J. (2003), “Compensating Wage Differentials for Fatal and Nonfatal Injury Risk by Gender and Race," *Journal of Risk and Uncertainty*, 27(3): 257–277.

Leicester, A. and Z. Oldfield (2009) "Using scanner technology to collect expenditure data" *Fiscal Studies*, vol. 30, no. 3/4, pp. 339–365

McFadden, D. (1974) "The Measurement of Urban Travel Demand," *Journal of Public Economics*, 3, 303-328

Manski, C.F. (2003), *Partial Identification of Probability Distributions*, Springer.

Nevo, Aviv (2001)“Measuring Market Power in the Ready-to-Eat Cereal Industry” *Econometrica*, Vol 69, No. 2 (March 2001), 307-342

Pakes, Ariel (2003), “A Reconsideration of Hedonic Price Indexes with an Application to PC’s," *The American Economic Review*, 93(5) 1578-1596.

Pesendorfer, Martin (2002) “Retail Sales: A Study of Pricing Behavior in Supermarkets” *Journal of Business*, Vol. 75, No. 1, January 2002

Pollak, Robert (1989), *The theory of the cost of living index*, Oxford University Press: New York.

Scotchmer, Susanne (1985), “Hedonic prices and cost/benefit analysis," *Journal of Economic Theory*, 37: 55-75.

Sattinger M. (1993), “Assignment Models of the Distribution of Earnings," *Journal of Economic Literature*, 31(2): 831–880.

Scotchmer, Susanne (1986), “The short run and long run benefits of environmental improvement," *Journal of Public Economics*, 30: 61-81.

Sieg, H., Smith, VK, Banzhaf, S., Walsh, R. (2004), “Estimating the General Equlibrium Benefits of Large Changes in Spatially Delineated Public Goods," *International Economic Review*, 45(4): 1047-69.

Smith, Howard (2004), “Supermarket Choice and Supermarket Competition in Market Equilibrium," *Review of Economic Studies*, 71(1): 235–263.

Smith, V. Kerry, and Huang, Ju-Chin (1995), “Can Markets Value Air Quality? A Meta-Analysis of Hedonic Property Value Models", *Journal of Political Economy*, 103(2): 209-27.

Train, Kenneth (2003) *Discrete Choice Methods with Simulation* Cambridge University Press

Triplett, J. (2003), “Using Scanner Data in Consumer Price Indexes: Some Neglected Conceptual Considerations”, in Scanner Data and Price Indexes.

Triplett, Jack (2005), “Handbook on hedonic indexes and quality adjustments in price indexes: special application to information technology products," *OECD Science, Technology and Industry Working Papers*, 2004/9, OECD Publishing.

Varian, Hal (1982), “The Nonparametric Approach to Demand Analysis," *Econometrica*, 50(4): 945–73.

Viscusi, W.K. and Aldy, J.E. (2003), “The value of a statistical life: a critical review of market estimates throughout the world," Journal of Risk and Uncertainty, 27(1): 5-76.

Table 1: Expenditure by food category

Product category	Observed purchases	Grossed up expenditure (£m)	Observed expenditure (£)	Share organic
Bread	642,298	761.8	501,195.50	1.9%
Biscuits	681,589	934.1	620,856.80	1.5%
Canned Goods				
Ambient Soup	145,824	155.7	101,279.60	1.3%
Baked Beans	160,385	151.1	103,654.40	0.7%
Canned Fruit	96,504	102.8	65,043.86	1.0%
Canned Milk Puddings	23,499	21.9	14,241.06	1.1%
Prepared Peas And Beans (M)	108,480	58.7	38,745.33	1.9%
Tomato Products	15,646	82.6	54,607.53	1.8%
Chilled Convenience				
Chilled Meat and Veg Extract	803	2.2	1,238.67	28.6%
Chilled Vegetarian Products	6,291	27.1	17,760.19	4.1%
Fresh Soup	708	2.2	1,278.55	11.7%
Other Chilled Convenience	26,504	73.2	49,201.10	0.6%
Chilled Drinks				
Chill One Shot Drinks (excl Flavoured Milk)	7,818	16.7	10,198.06	1.0%
Chilled Fruit Juices	90,977	215.9	134,044.50	5.1%
Dairy Products				
Butter	110,590	223.1	140,477.50	1.7%
Cream	97,275	124.9	77,871.22	1.4%
Hens Eggs	195,032	322.1	207,443.00	5.8%
Desserts Long Life	18,992	22.9	16,063.18	2.2%
Cheese (excl Formage Frais), pre-packaged	313,840	687.6	454,997.00	0.6%
Cheese (excl Formage Frais), loose	199,531	547.0	354,582.90	1.4%
Milk	745,006	1460.8	952,108.50	2.2%
Yoghurt	384,261	581.7	373,848.80	5.7%
Chilled/Prepared Fruit and Veg	150,201	284.6	178,726.10	1.5%
Frozen Prepared Foods				
Frozen Vegetables	95,256	160.5	107,926.80	2.6%
Ice Cream	74,719	195.3	125,001.00	1.3%
Fruit & Vegetables				
Fruit, pre-packaged	112,701	323.3	196,801.60	1.1%
Fruit, loose	1,003,804	1681.9	1,057,302.00	2.5%
Vegetables (brassicas, legumes, root), pre-packaged	240,272	459.1	301,242.10	2.6%
Vegetables (brassicas, legumes, root), loose	369,969	438.3	275,008.20	2.2%
Vegetables (brassicas, legumes, root), other	436,134	364.2	232,421.90	5.1%
Vegetables (salad and other), loose	814,249	941.9	588,532.40	2.7%
Vegetables (salad and other), pre-packaged	16,615	28.7	17,886.60	0.7%
Hot Beverages				
Food Drinks	43,780	93.9	59,194.17	1.6%
Fruit And Herbal Teas, pre-packaged	10,784	18.9	11,463.78	3.4%
Fruit And Herbal Teas, loose	2,787	4.9	3,192.41	6.4%

Table 1: Expenditure by food category

Product category	Observed purchases	Grossed up expenditure (£m)	Observed expenditure (£)	Share organic
Instant Coffee	94,000	380.2	243,416.40	1.4%
Coffee (Beans, Ground or Liquid)	16,020	65.3	37,546.37	12.5%
Tea	103,048	290.4	188,799.80	1.4%
Meat				
Fresh Bacon Rashers	153,469	516.8	337,446.00	0.4%
Fresh Beef/Veal	180,752	843.0	552,037.10	1.1%
Fresh Lamb	36,127	244.9	155,213.50	2.0%
Fresh Pork	83,703	383.0	250,903.50	0.7%
Fresh Sausages	95,486	221.4	147,070.00	0.8%
Packet & Other Foods				
Breakfast Cereals, pre-packaged	303,306	760.5	505,664.40	1.4%
Breakfast Cereals, loose	43,461	93.7	62,915.65	1.8%
Cous Cous	6,390	8.1	5,188.06	4.6%
Dry Meat Substitutes	1,232	2.0	1,352.80	8.1%
Dry Pasta	82,424	79.1	52,887.59	2.5%
Flour	46,616	47.2	29,392.26	4.8%
Home Baking	77,243	147.8	91,268.00	2.1%
Honey	11,382	36.2	21,843.02	4.4%
Lemon And Lime Juices	6,981	7.1	4,303.94	1.2%
Peanut Butter	11,980	19.5	12,690.15	1.4%
Preserves	71,325	98.6	61,326.38	3.0%
Sugar	116,556	163.9	107,187.90	0.5%
Syrup And Treacle	6,071	8.4	5,364.67	3.4%
Vinegar	18,928	20.2	12,595.02	3.6%
Condiments				
Ambient Condiments	7,335	9.8	6,337.51	3.0%
Ambient Salad Accompaniments	26,348	39.5	25,285.99	1.8%
Pickles	31,026	42.7	27,267.50	0.6%
Sauces And Ketchup (Exc:Wrcster)	71,970	105.5	72,759.39	1.6%
Fresh Poultry	63,472	370.8	242,645.70	2.1%
Savouries				
Nuts	60,107	122.1	76,189.25	1.2%
Savoury Snacks And Reconstitutes	200,674	321.3	227,365.20	2.1%
Ambient Savoury Snacks	123,438	163.9	111,628.90	0.6%
Savoury Home Cooking				
Ambient Cooking Scauces excl Condiments	227,456	375.4	251,363.50	1.7%
Ambient Meat+Veg Extracts	98,116	147.6	98,092.15	1.3%
Cooking Oils	53,188	119.3	77,575.33	1.7%
Soft Drinks				
Bottled Non-Lemonade (flavoured)	130,436	213.8	146,018.10	6.1%
Canned Lemonade	479	0.7	502.07	21.5%
Canned Non-Lemonade (flavoured)	29,837	77.8	52,673.88	1.0%

Table 1: Expenditure by food category

Product category	Observed purchases	Grossed up expenditure (£m)	Observed expenditure (£)	Share organic
Ambient Flavoured Milk	10,328	17.2	11,836.07	6.9%
Ambient Fruit Juices	173,282	294.5	189,479.30	1.5%
Ambient One Shot Drinks	80,601	159.5	114,693.10	1.2%
Chocolate and Sugar Confectionary	540,699	1130.6	740,970.10	2.4%
Total	11,100,000	19721.3	12,800,000.00	2.1%

Note: Data include 16,881 households over calendar year 2004. A purchase is a household-store-day transaction (if a household buys two of the exact same product in one day at the same store this is one transaction, if they buy the same product at a different store or a different product at the same store that counts as a separate transaction). Grossed-up expenditure is sampled expenditure weighted by household demographic weight in the sample relative to the UK population.

Table 2: Share of household expenditure on products with organic characteristic

Share of household expenditure on organic products	**Number of households**	**% of households**	**Cumulative % of households**
0	3,177	18.82	18.82
less than 0.25%	2,168	12.84	31.66
0.25% - 0.5%	2,279	13.50	45.16
0.5% - 1%	3,007	17.81	62.98
1% - 5%	4,985	29.53	92.51
5% - 10%	757	4.48	96.99
over 10%	508	3.01	100.00
Total	16,881	100.00	

Note: Data include 16,881 households over calendar year 2004.

Table 3: Hedonic regressions for milk

Dependent variable: ln(price)	(1)	(2)	(3)	(4)
organic	0.296***	0.256***	0.150***	0.166***
	[0.0708]	[0.0381]	[0.0225]	[0.0385]
organic*semi-skimmed	-0.0837			
	[0.0832]			
organic*skimmed	-0.031			
	[0.0955]			
organic*Asda				-0.170*
				[0.0919]
organic*M&S				0.061
				[0.0385]
organic*Safeway				0.0302
				[0.0354]
organic*Sainsbury				0.0298
				[0.0369]
organic*Tesco				-0.0343
				[0.0349]
organic*Waitrose				0.136***
				[0.0506]
fat: semi-skimmed	-0.0542	-0.0564	0.00278	0.00279
	[0.0476]	[0.0464]	[0.00541]	[0.00538]
fat: skimmed	-0.145**	-0.147**	-0.0232**	-0.0234**
	[0.0616]	[0.0602]	[0.00962]	[0.00957]
size (1.136 litre, 2 pints)			-0.0714***	-0.0717***
			[0.00942]	[0.00943]
size (1.14-2.272 litres, inc. 4 pints)			-0.191***	-0.191***
			[0.0100]	[0.0100]
size (3 litres and above)			-0.215***	-0.216***
			[0.0104]	[0.0104]
brand: budget private label			-0.395***	-0.393***
			[0.0598]	[0.0598]
brand: standard private label			-0.0823***	-0.0802***
			[0.0194]	[0.0192]
container: carton			-0.231***	-0.235***
			[0.0459]	[0.0473]
container: other			-0.314***	-0.314***
			[0.0493]	[0.0493]
container: plastic			-0.290***	-0.292***
			[0.0203]	[0.0201]
type: Channel Island			-0.527***	-0.531***
			[0.163]	[0.165]
type: Ordinary			-0.940***	-0.946***
			[0.156]	[0.158]

Table 3: Hedonic regressions for milk

Dependent variable: ln(price)	(1)	(2)	(3)	(4)
type: other non-cows			-0.289***	-0.289***
			[0.0979]	[0.0999]
type: Soya			-0.612***	-0.609***
			[0.170]	[0.171]
type: other			-0.389**	-0.394**
			[0.152]	[0.154]
treatment: pasteurised			-0.136***	-0.138***
			[0.0198]	[0.0196]
treatment: sterilised			-0.170***	-0.170***
			[0.0407]	[0.0403]
treatment: U.H.T.			-0.316***	-0.315***
			[0.0441]	[0.0450]
treatment: other types			-0.142	-0.145
			[0.121]	[0.122]
Asda			-0.0192**	-0.0165*
			[0.00929]	[0.00928]
MandS			0.0586***	0.0541***
			[0.00980]	[0.00966]
Safeway			0.00654	0.0048
			[0.00836]	[0.00841]
Sainsbury			-0.00547	-0.00786
			[0.0102]	[0.0104]
Tesco			-0.0155	-0.0155
			[0.00954]	[0.00942]
Waitrose			0.00768	-0.0103
			[0.0157]	[0.0120]
ticket price reducation			-0.0611***	-0.0631***
			[0.0195]	[0.0196]
multi-purchase			-0.392***	-0.391***
			[0.0235]	[0.0233]
extra free			-0.0105	-0.00915
			[0.0680]	[0.0675]
Constant	-0.646***	-0.644***	0.851***	0.858***
	[0.0336]	[0.0330]	[0.157]	[0.159]
Adjusted R-squared	0.065	0.065	0.726	0.728

*Note: Regression include 745,006 observations on16,881 households purchases of milk over calendar year 2004; all regressions include month and region effects. Standard errors in [] are clustered at the barcode level and allow for general correlation. *** p<0.01, ** p<0.05, * p<0.1.The omitted characteristics are other stores, full-fat,1 pint or smaller, brand name,bottled, buttermilk and filtered.*

Table 4: Organic coefficients

Product category	organic		organic*Asda		organic*Sainsbury		organic*Tesco		organic*Waitrose		adjusted R^2	N
Bread	0.177	*(0.031)*	-0.060	*(0.059)*	0.056	*(0.086)*	0.043	*(0.053)*	-0.068	*(0.106)*	0.818	642,298
Biscuits	-0.180	*(0.160)*	0.147	*(0.201)*	0.455	*(0.171)*	0.315	*(0.199)*	0.331	*(0.168)*	0.730	681,589
Canned Goods												
Ambient Soup	0.119	*(0.082)*	-0.112	*(0.079)*	0.368	*(0.170)*	0.023	*(0.034)*	-0.258	*(0.075)*	0.755	145,824
Baked Beans	0.519	*(0.126)*	-0.192	*(0.127)*	-0.034	*(0.176)*	-0.233	*(0.129)*	-0.191	*(0.272)*	0.943	160,385
Canned Fruit	0.561	*(0.176)*	-0.198	*(0.180)*			0.077	*(0.111)*			0.625	96,504
Canned Milk Puddings	0.327	*(0.060)*	-0.192	*(0.082)*			-0.155	*(0.113)*			0.768	23,499
Prepared Peas And Beans	0.531	*(0.093)*	0.059	*(0.121)*	0.073	*(0.125)*	-0.006	*(0.095)*	-0.070	*(0.119)*	0.766	108,480
Tomato Products	0.293	*(0.370)*	0.000	*(0.000)*	-0.018	*(0.392)*	0.000	*(0.000)*	0.622	*(0.375)*	0.692	15,646
Chilled Convenience												
Chilled Meat and Veg Extract	1.249	*(0.144)*			-0.611	*(0.091)*			-0.166	*(0.111)*	0.788	803
Chilled Vegetarian Products	0.000	*(0.000)*			0.000	*(0.000)*	0.000	*(0.000)*	0.000	*(0.000)*	0.852	6,291
Fresh Soup	0.162	*(0.137)*					0.113	*(0.034)*			0.742	708
Other Chilled Convenience	-1.047	*(0.505)*					0.001	*(0.045)*	-0.118	*(0.083)*	0.799	26,504
Chilled Drinks												
Chill One Shot Drinks (excl Flavoured Milk)	1.038	*(0.097)*									0.870	7,818
Chilled Fruit Juices	0.184	*(0.171)*	0.052	*(0.187)*	0.115	*(0.213)*	0.084	*(0.181)*	0.027	*(0.199)*	0.710	90,977
Dairy Products												
Butter	0.314	*(0.069)*	-0.054	*(0.082)*	-0.046	*(0.080)*	-0.291	*(0.094)*	-0.069	*(0.063)*	0.723	110,590
Cream	-0.070	*(0.073)*	-0.157	*(0.051)*	0.035	*(0.022)*	-0.070	*(0.083)*	0.077	*(0.048)*	0.794	97,275
Hens Eggs	0.310	*(0.054)*	-0.051	*(0.057)*	0.183	*(0.063)*	0.133	*(0.064)*	0.130	*(0.095)*	0.781	195,032
Desserts Long Life	-0.038	*(0.024)*	0.046	*(0.027)*	0.029	*(0.042)*	-0.024	*(0.029)*			0.496	18,992
Cheese (excl Formage Frais), pre-packaged	0.015	*(0.089)*	0.291	*(0.142)*	0.167	*(0.128)*	0.216	*(0.147)*	0.148	*(0.111)*	0.765	313,840
Cheese (excl Formage Frais), loose	0.340	*(0.077)*	-0.191	*(0.096)*	-0.058	*(0.086)*	-0.136	*(0.098)*	-0.259	*(0.110)*	0.627	199,531
Milk	0.167	*(0.039)*	-0.169	*(0.093)*	0.030	*(0.037)*	-0.038	*(0.035)*	0.134	*(0.051)*	0.727	745,006
Yoghurt	0.212	*(0.042)*	-0.013	*(0.064)*	-0.057	*(0.032)*	-0.010	*(0.026)*	-0.184	*(0.035)*	0.680	384,261

Table 4: Organic coefficients

Product category	organic		organic*Asda		organic*Sainsbury		organic*Tesco		organic*Waitrose		adjusted R^2	N
Chilled/Prepared Fruit and Veg	-0.003	*(0.070)*			0.404	*(0.104)*	0.362	*(0.181)*	0.208	*(0.213)*	0.719	150,201
Frozen Prepared Foods												
Frozen Vegetables	0.176	*(0.067)*			0.054	*(0.068)*	0.165	*(0.076)*	-0.141	*(0.096)*	0.644	95,256
Ice Cream	0.045	*(0.067)*			0.320	*(0.091)*	0.287	*(0.087)*	0.068	*(0.111)*	0.874	74,719
Fruit & Vegetables												
Fruit, pre-packaged	0.165	*(0.150)*	0.111	*(0.170)*	0.067	*(0.224)*	-0.014	*(0.198)*	-0.106	*(0.196)*	0.859	112,701
Fruit, loose	0.174	*(0.145)*	-0.022	*(0.167)*	0.351	*(0.153)*	0.066	*(0.163)*	0.113	*(0.178)*	0.781	1,003,804
Vegetables (brassicas, legumes, potatoes, root crops), pre-packaged	0.101	*(0.091)*	-0.186	*(0.114)*	0.215	*(0.132)*	0.030	*(0.128)*	0.035	*(0.162)*	0.653	240,272
Vegetables (brassicas, legumes, potatoes, root crops), loose	0.216	*(0.103)*	0.286	*(0.151)*	0.115	*(0.110)*	0.205	*(0.126)*	0.069	*(0.137)*	0.806	369,969
Vegetables (brassicas, legumes, potatoes, root crops), other	0.536	*(0.091)*	-0.078	*(0.114)*	-0.152	*(0.114)*	-0.087	*(0.104)*	0.033	*(0.117)*	0.753	436,134
Vegetables (salad and other), loose	0.461	*(0.073)*	-0.056	*(0.107)*	0.109	*(0.106)*	-0.004	*(0.088)*	-0.023	*(0.108)*	0.742	814,249
Vegetables (salad and other), pre-packaged	-0.542	*(0.601)*					0.737	*(0.643)*			0.920	16,615
Hot Beverages												
Food Drinks	0.625	*(0.057)*			-0.017	*(0.040)*	0.024	*(0.081)*	-0.094	*(0.067)*	0.876	43,780
Fruit And Herbal Teas, pre-packaged	0.427	*(0.094)*			-0.048	*(0.065)*	-0.393	*(0.122)*			0.654	10,784
Fruit And Herbal Teas, loose	0.148	*(0.071)*	-0.063	*(0.060)*			-0.069	*(0.078)*			0.841	2,787
Instant Coffee	0.397	*(0.086)*	-0.090	*(0.080)*	-0.167	*(0.045)*	-0.130	*(0.055)*	0.040	*(0.051)*	0.731	94,000
Coffee (Beans, Ground or Liquid)	0.083	*(0.132)*	-0.157	*(0.141)*	-0.078	*(0.130)*	-0.050	*(0.150)*	-0.192	*(0.122)*	0.710	16,020
Tea	0.486	*(0.087)*	-0.292	*(0.128)*	-0.107	*(0.074)*	-0.067	*(0.124)*	-0.388	*(0.084)*	0.864	103,048
Meat												
Fresh Bacon Rashers	0.696	*(0.245)*	-0.646	*(0.244)*	-0.021	*(0.260)*	-0.273	*(0.248)*	-0.045	*(0.240)*	0.590	153,469
Fresh Beef/Veal	0.378	*(0.082)*	-0.052	*(0.097)*	0.140	*(0.122)*	0.114	*(0.182)*	-0.086	*(0.121)*	0.742	180,752
Fresh Lamb	0.144	*(0.084)*	-0.059	*(0.099)*	0.132	*(0.101)*	0.048	*(0.092)*	0.034	*(0.156)*	0.601	36,127
Fresh Pork	0.210	*(0.178)*	0.022	*(0.188)*	0.420	*(0.202)*	0.297	*(0.184)*	0.424	*(0.201)*	0.560	83,703
Fresh Sausages	0.400	*(0.093)*			0.052	*(0.107)*	0.255	*(0.088)*	0.010	*(0.117)*	0.767	95,486
Packet & Other Foods												

Table 4: Organic coefficients

Product category	organic		organic*Asda		organic*Sainsbury		organic*Tesco		organic*Waitrose		adjusted R^2	N
Breakfast Cereals, pre-packaged	0.112	*(0.067)*	-0.084	*(0.067)*	0.102	*(0.064)*	0.172	*(0.065)*	0.116	*(0.071)*	0.851	303,306
Breakfast Cereals, loose	0.307	*(0.044)*	0.066	*(0.048)*	0.060	*(0.048)*	0.048	*(0.045)*	0.003	*(0.026)*	0.860	43,461
Cous Cous	0.171	*(0.083)*			0.083	*(0.049)*	0.103	*(0.054)*			0.937	6,390
Dry Meat Substitutes	0.249	*(0.080)*					0.177	*(0.073)*			0.776	1,232
Dry Pasta	0.219	*(0.132)*	-0.244	*(0.229)*	0.191	*(0.164)*	0.148	*(0.096)*	0.511	*(0.171)*	0.811	82,424
Flour	0.052	*(0.095)*	-0.193	*(0.105)*	-0.128	*(0.059)*	0.013	*(0.102)*	0.346	*(0.127)*	0.904	46,616
Home Baking	0.226	*(0.115)*	-0.086	*(0.163)*	-0.231	*(0.130)*	-0.071	*(0.132)*	-0.033	*(0.167)*	0.763	77,243
Honey	0.192	*(0.052)*	-0.035	*(0.039)*	0.038	*(0.065)*	0.173	*(0.064)*	-0.068	*(0.051)*	0.648	11,382
Lemon And Lime Juices	0.579	*(0.058)*			-0.137	*(0.061)*					0.875	6,981
Peanut Butter	0.450	*(0.058)*					0.193	*(0.069)*			0.798	11,980
Preserves	0.182	*(0.097)*	0.132	*(0.101)*	0.286	*(0.108)*	0.001	*(0.090)*	0.297	*(0.120)*	0.812	71,325
Sugar	0.455	*(0.143)*	-0.213	*(0.068)*	-0.277	*(0.119)*	-0.013	*(0.057)*	-0.098	*(0.061)*	0.813	116,556
Syrup And Treacle	1.143	*(0.077)*			0.287	*(0.091)*					0.703	6,071
Vinegar	0.274	*(0.220)*			-0.165	*(0.182)*	-0.146	*(0.117)*	-0.235	*(0.130)*	0.852	18,928
Condiments												
Ambient Condiments	0.609	*(0.153)*			-0.010	*(0.043)*	0.047	*(0.029)*			0.485	7,335
Ambient Salad Accompaniments	0.852	*(0.084)*			-0.028	*(0.089)*	0.013	*(0.127)*	-0.020	*(0.084)*	0.899	26,348
Pickles	0.440	*(0.110)*					-0.195	*(0.083)*			0.861	31,026
Sauces And Ketchup (Exc:Wrcster)	0.101	*(0.082)*	0.059	*(0.039)*	0.067	*(0.093)*	0.073	*(0.037)*	0.248	*(0.155)*	0.840	71,970
Fresh Poultry	0.492	*(0.096)*	0.187	*(0.147)*	0.062	*(0.143)*	0.309	*(0.214)*	-0.533	*(0.184)*	0.562	63,472
Savouries												
Nuts	0.436	*(0.130)*			-0.200	*(0.207)*	-0.047	*(0.201)*	-0.010	*(0.194)*	0.460	60,107
Savoury Snacks And Reconstitutes	0.133	*(0.164)*	0.140	*(0.143)*	0.192	*(0.158)*	0.164	*(0.146)*	0.113	*(0.142)*	0.664	200,674
Ambient Savoury Snacks	-0.181	*(0.106)*	0.200	*(0.100)*	0.343	*(0.093)*	0.141	*(0.222)*	0.290	*(0.121)*	0.833	123,438
Savoury Home Cooking												
Ambient Cooking Scauces excl Condiments	0.139	*(0.061)*	-0.192	*(0.054)*	0.047	*(0.058)*	-0.028	*(0.065)*	-0.017	*(0.065)*	0.876	227,456
Ambient Meat+Veg Extracts	0.381	*(0.143)*	-0.235	*(0.082)*	0.056	*(0.086)*	-0.236	*(0.070)*	0.058	*(0.099)*	0.723	98,116
Cooking Oils	0.072	*(0.188)*	0.276	*(0.152)*	0.358	*(0.190)*	0.537	*(0.283)*	0.307	*(0.231)*	0.853	53,188

Table 4: Organic coefficients

Product category	organic		organic*Asda		organic*Sainsbury		organic*Tesco		organic*Waitrose		adjusted R^2	N
Soft Drinks												
Bottled Non-Lemonade (flavoured)	0.810	*(0.189)*	0.047	*(0.085)*	-0.139	*(0.063)*	-0.101	*(0.041)*	-0.305	*(0.068)*	0.867	130,436
Canned Lemonade	0.631	*(0.223)*					-0.303	*(0.207)*			0.953	479
Canned Non-Lemonade (flavoured)	-0.074	*(0.131)*					0.521	*(0.144)*	-0.414	*(0.132)*	0.783	29,837
Ambient Flavoured Milk	0.062	*(0.162)*	0.130	*(0.105)*	-0.106	*(0.106)*	-0.017	*(0.160)*			0.761	10,328
Ambient Fruit Juices	0.415	*(0.074)*	-0.093	*(0.063)*	0.038	*(0.071)*	0.009	*(0.059)*	0.117	*(0.081)*	0.784	173,282
Ambient One Shot Drinks	0.311	*(0.073)*	0.250	*(0.085)*	0.025	*(0.080)*	0.114	*(0.082)*			0.666	80,601
Chocolate and Sugar Confectionary	0.276	*(0.054)*	-0.201	*(0.049)*	0.022	*(0.061)*	-0.051	*(0.050)*	0.198	*(0.067)*	0.642	540,699

Notes: Each row represents a separate hedonic regression. An observation is a transaction. The coefficient and standard error are those on a dummy for whether the specific product (bar code) is organic. Standard errors are clustered at the product level. The adjusted R^2 is from the overall regression. The numbers indicate the number of characteristics of each type appear in the hedonic regression. For example, for Bacon Rashers there are 3 brand characteristics (Branded, Budget or Standard Private Label), 4 origin characteristics (Britain, Ireland, Northern Europe and Other) and 2 variety characteristics (smoked or unsmoked). In all regressions there are 5 size categories, 8 store indicators (see Table 2), time and region effects. The final column shows the number of observations.

Table 5: estimated household lower bound on willingness to pay for organic over a year

Household lower bound on willingness to pay for organic over a year	Number of households	% of households	Cumulative % of households
0	3,918	23.21	23.21
less than £1	4,401	26.07	49.28
£1 - £5	5,361	31.76	81.04
£5 - £10	1,540	9.12	90.16
£10 - £50	1,396	8.27	98.43
£50 - £100	142	0.84	99.27
over £100	123	0.73	100.00
Total	16,881	100.00	

Note: Data include 16,881 households over calendar year 2004.Each household's lower bound is calculated as in equation (4) using the estimated coefficients summarised in Table 4.

Table 6: estimated household lower bound on willingness to pay as a share of expenditure on organic products

Household lower bound on willingness to pay as a share of expenditure on organic products	Number of households	% of households	Cumulative % of households
0	3,918	23.21	23.21
less than 10%	1,197	7.09	30.30
10% - 15%	1,779	10.54	40.84
15% - 20%	2,709	16.05	56.89
20% - 25%	2,873	17.02	73.91
25% - 50%	4,114	24.37	98.28
over 50%	291	1.72	100.00
Total	16,881	100.00	

Note: Data include 16,881 households over calendar year 2004.Each household's lower bound is calculated as in equation (4), using the estimated coefficients summarised in Table 4, and taken as a share of the household's total expenditure on organic products.

Table 7: estimated household upper bound on willingness to pay as a share of expenditure on non-organic products

Household upper bound on willingness to pay as a share of expenditure on organic products	Number of households	% of households	Cumulative % of households
0	64	0.38	0.38
less than 20%	498	2.95	3.33
20% - 30%	5,738	33.99	37.32
30% - 40%	8,615	51.03	88.35
40% - 50%	1,869	11.07	99.43
over 50%	97	0.57	100.00
Total	16,881	100.00	

Note: Data include 16,881 households over calendar year 2004. Each household's upper bound is calculated as in equation (5), using the estimated coefficients summarised in Table 4, and taken as a share of the household's total expenditure on non-organic products.

Table 8: Health, share of expenditure on organic, lower bound on wtp as share of total expenditure and number of households

		Organic Products Are Healthier	
		Agree Strongly/Agree	Agree Nor Disagree/Disagree /Strongly Disagree
I Try To Buy a Healthy Range Of Foods These Days	Agree Strongly/Agree	4.7% 2.2% (2178)	1.8% 0.4% (512)
	Agree Nor Disagree/Disagree /Strongly Disagree	1.2% 0.3% (4150)	1.3% 0.3% (6649)

Note: In each cell the top % indicates the share of total household expenditure that is on organic products, the second % indicates the mean lower bound on willingness to pay by households in that cell, and the number in () indicates the number of households that selected the indicated response.

Table 9: Environment, share of expenditure on organic, lower bound on wtp as share of total expenditure and number of households

		Organic Products Are Friendlier To The Environment	
		Agree Strongly/Agree	Agree Nor Disagree/Disagree /Strongly Disagree
I Try To Buy Environmentally Friendly Products	Agree Strongly/Agree	3.7% 0.9% (1941)	2.9% 0.7% (1723)
	Agree Nor Disagree/Disagree /Strongly Disagree	1.2% 0.3% (2087)	1.3% 0.3% (7336)

Note: In each cell the top % indicates the share of total household expenditure that is on organic products, the second % indicates the mean lower bound on willingness to pay by households in that cell, and the number in () indicates the number of households that selected the indicated response.

Table 10: Quality, share of expenditure on organic and number of households

`	Organic Foods Are Better Quality

		Agree Strongly/Agree	Agree Nor Disagree/Disagree /Strongly Disagree
I Don't Mind Paying For Quality	Agree Strongly/Agree	5.9% 1.4% (1334)	2.6% 0.6% (418)
	Agree Nor Disagree/Disagree /Strongly Disagree	1.4% 0.3% (4712)	1.4% 0.3% (6706)

Note: In each cell the top % indicates the share of total household expenditure that is on organic products, the second % indicates the mean lower bound on willingness to pay by households in that cell, and the number in () indicates the number of households that selected the indicated response.

Table 11: Determinants of lower bound on willingness to pay for organic

Dep var: lower bound on willingness to pay for organic in £	(1)	(2)	(3)	(4)	(5)	(6)
	All households	All households	All households	Household Professional Class (AB)	Household Skilled Class (C)	Household Unskilled Class (DE)
Health important	5.783***	5.307***	4.784***	7.676***	4.641***	3.914***
	[0.534]	[0.549]	[0.515]	[1.989]	[0.658]	[0.864]
Health (upper right quadrant of Table 8)		-0.962	-1.147*	-5.747*	-1.007	-0.125
		[0.601]	[0.610]	[3.126]	[0.858]	[0.707]
Health (bottom left quadrant of Table 8)		-0.608***	-1.026***	-2.382**	-1.252***	-0.102
		[0.154]	[0.184]	[1.129]	[0.220]	[0.182]
Environment important	1.733***	2.195***	1.689***	2.277	1.581**	1.822**
	[0.651]	[0.566]	[0.524]	[2.317]	[0.643]	[0.790]
Environment (upper right quadrant of Table 9)		0.727	0.414	3.98*	0.249	-0.131
		[0.466]	[0.481]	[2.302]	[0.615]	[0.562]
Environment (bottom left quadrant of Table 9)		-0.0115	-0.36*	-1.766*	-0.114	-0.338
		[0.172]	[0.187]	[1.021]	[0.243]	[0.236]
Quality important	8.711***	8.522***	7.988***	8.239**	9.126***	4.955***
	[1.129]	[1.159]	[1.095]	[3.906]	[1.421]	[1.434]
Quality (upper right quadrant of Table 10)		-0.647	-0.341	1.882	-0.873	0.806
		[0.839]	[0.799]	[4.642]	[0.798]	[1.519]
Quality (bottom left quadrant of Table 10)		-0.102	-0.736***	-2.669**	-0.461**	-0.597*
		[0.179]	[0.213]	[1.208]	[0.227]	[0.327]
Household Class A or B			3.126***			
			[0.799]			
Household class C1 or C2			0.346			
			[0.267]			
Single young			3.06***	7.009	2.854***	2.156***
			[0.747]	[5.692]	[0.907]	[0.568]

Table 11: Determinants of lower bound on willingness to pay for organic

Dep var: lower bound on willingness to pay for organic in £	(1)	(2)	(3)	(4)	(5)	(6)
	All households	All households	All households	Household Professional Class (AB)	Household Skilled Class (C)	Household Unskilled Class (DE)
Single with kids			1.799***	9.439	0.979	1.707***
			[0.570]	[6.832]	[0.712]	[0.522]
Single pensioner			3.064***	6.426	1.757*	2.974***
			[0.732]	[5.988]	[0.989]	[0.612]
Couple no kids			2.389***	3.669	2.286***	1.974***
			[0.569]	[3.558]	[0.728]	[0.735]
Couple with kids			0.729	-0.325	0.867	0.246
			[0.456]	[2.664]	[0.661]	[0.433]
Couple pensioner			1.697***	1.887	2.016***	1.511***
			[0.527]	[3.228]	[0.920]	[0.452]
Other no kids			1.796***	0.653	1.116	3.003***
			[0.599]	[3.289]	[0.738]	[0.946]
At least one vegetarian in the household			3.807***	3.853	5.061**	0.5
			[1.620]	[4.957]	[2.284]	[1.061]
Annual expenditure on alcohol, food, toiletries and cleaning products (in £,000s)			3.21***	8.19**	2.76***	2.21***
			[0.526]	[3.96]	[0.421]	[0.299]
Response to attitudinal question missing	-1.143***	-1.07***	-0.89***	-3.139*	-0.706*	-0.839***
	[0.212]	[0.217]	[0.248]	[1.759]	[0.384]	[0.261]
Constant	2.204***	2.417***	-5.65***	-13.54	-4.391***	-3.559***
	[0.112]	[0.156]	[1.223]	[9.862]	[1.182]	[0.707]
Observations	13591	13591	13489	1343	7781	4365
R-squared	0.058	0.059	0.091	0.107	0.092	0.11

*Notes: Standard errors in [] are robust. See Table 12 for means of variables and notes to Table 12 for definition of social class. *** p<0.01, ** p<0.05, * p<0.1*

Table 12: mean of demographic variables

Variable	mean (s.d.)
Family type = Single young	0.082
	(0.274)
Family type =Single parent	0.062
	(0.242)
Family type = Single pensioner	0.075
	(0.263)
Family type = Couple no children	0.135
	(0.341)
Family type = Couple with children	0.396
	(0.489)
Family type = Pensioner couple	0.117
	(0.321)
Family type = Others no children	0.085
	(0.278)
Family type = Others with children	0.0483
	(0.137)
Annual expenditure on alcohol, food, toiletries and cleaning products	2048.07
	(1048.23)
Household Professional Class (A or B)	0.100
	(0.300)
Household Skilled Class (C1 or C2)	0.577
	(0.494)
At least one vegetarian in the household	0.023
	(0.149)
Demographics or attitudes missing	0.036
	(0.187)

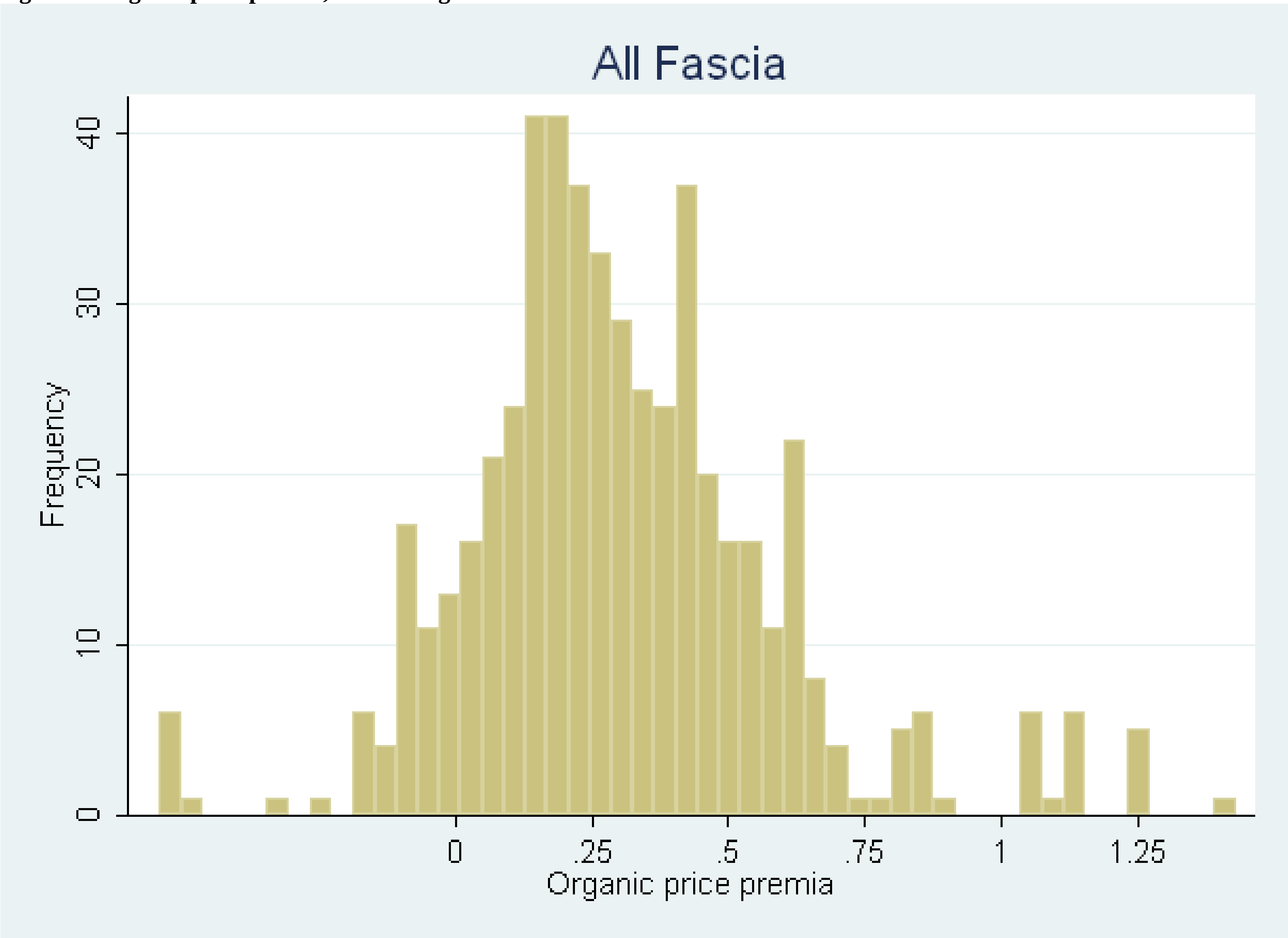
Notes: Social class is A (upper middle class - higher managerial, administrative or professional), B (middle class - intermediate managerial, administrative or professional) C1 (lower middle class - supervisory or clerical, junior managerial, administrative or professional) or C2 (skilled working class - skilled manual workers) (the omitted category is D (working class - semi and unskilled manual workers) and E (those at lowest level of subsistence - state pensioners or widows (no other earner), casual or lowest grade workers).

Table 13: Upper bound on revenue increase from converting to only organic eggs

Fascia	Price Premia on organic	Expenditure (£m)	Share of expenditure on organic	Expenditure if all purchases were organic (£m)	Upper bound on % increase in revenue if all products organic
Eggs					
Asda	26%	54.4	4.0%	69.8	28.4%
M&S	42%	3.0	19.8%	4.3	42.0%
Sainsbury	40%	50.2	14.8%	77.3	54.2%
Safeway	49%	20.0	5.9%	29.2	46.4%
Tesco	44%	94.3	6.2%	143.2	52.1%
Waitrose	44%	7.3	13.6%	10.7	47.6%
Other	31%	93.0	2.9%	125.3	35.1%
Vegetables (salad and other), loose					
Asda	40%	139.4	1.7%	207.4	48.9%
M&S	15%	18.1	4.6%	20.9	15.1%
Sainsbury	57%	182.4	5.0%	314.5	72.7%
Safeway	47%	62.4	1.0%	99.0	58.9%
Tesco	46%	280.0	3.3%	436.2	55.9%
Waitrose	44%	31.6	15.3%	46.2	46.5%
Other	46%	228.1	0.8%	359.0	57.8%

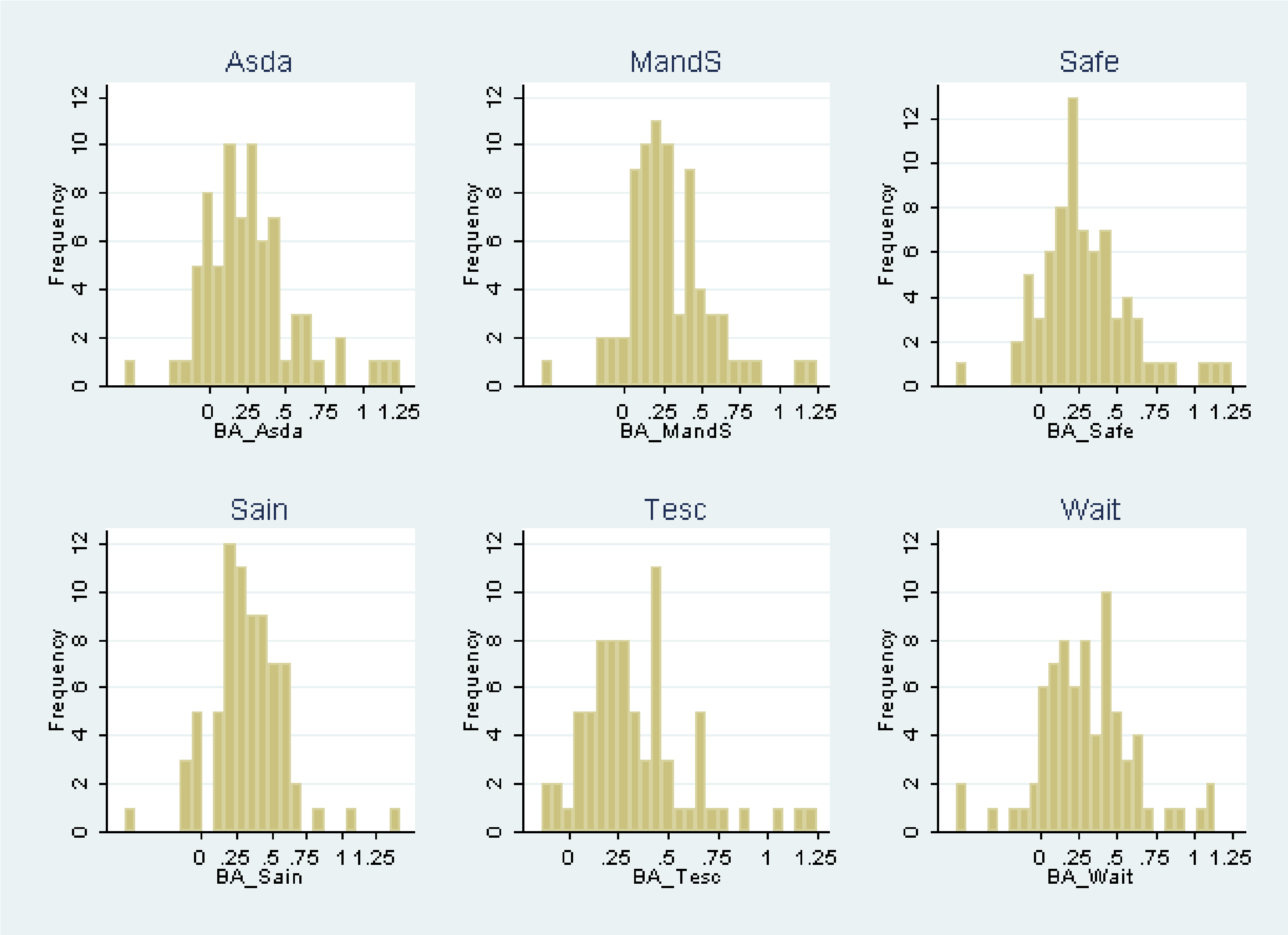
Note: Data include 16,881 households over calendar year 2004. All values are sampled expenditure grossed-up by household demographic weight in the sample relative to the UK population.

Figure 1: Organic price premia, across all goods and store fascia



Notes: The histogram shows the (unweighted) distribution of all 518 estimated price premia on the organic characteristic.

Figure 2: Organic price premia, by store fascia (only coefficients that are significantly different than zero at 5% level)



Notes: The histogram shows the (unweighted) distribution of all 74 estimated price premia on the organic characteristic by fascia.