Economics Letters 120 (2013) 284–287

Contents lists available at SciVerse ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

Hedonic methods for baskets of goods

Rachel Griffith [a,b], Lars Nesheim [c,d,*]

[a] Institute for Fiscal Studies, United Kingdom
[b] University of Manchester, United Kingdom
[c] CeMMAP, Institute for Fiscal Studies, United Kingdom
[d] UCL, United Kingdom

HIGHLIGHTS

- We apply hedonic price methods to large complex baskets of goods.
- We combine hedonic price methods with revealed preference.
- We estimate bounds on willingness-to-pay for organic using scanner data with thousands of goods.
- Median lower bound is 0.2% of expenditure but 12.5% have a lower bound larger than 1%.
- Median upper bound is 31.5% with most households having a bound between 20% and 40%.

ARTICLE INFO

Article history:
Received 28 February 2013
Accepted 24 April 2013
Available online 3 May 2013

JEL classification:
C8
D1
D4
L1
L8

Keywords:
Hedonic prices
Scanner data
Willingness to pay
Organic

ABSTRACT

Existing hedonic methods cannot be easily adapted to estimate willingness to pay for product characteristics when willingness to pay depends on a very large basket of goods. We show how to marry these methods with revealed preference arguments to estimate bounds on willingness to pay using data on purchases of seemingly impossibly high dimensional baskets of goods. 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.

1. Introduction

Researchers, policy makers and firms often want to estimate consumers' willingness to pay for a characteristic of a good. For example, there is much interest in estimating willingness to pay for organic products (for example, see Blow et al., 2008). For small scale problems, hedonic or discrete choice methods can provide estimates. However, these methods are not tractable when the number of relevant products is large or the characteristic space is large. We consider the (common) situation in which a consumer buys a large basket of goods, each good having many characteristics. We propose a method to marry hedonic price methods to revealed preference methods for analysing these large and complex baskets of goods.

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 and willingness to accept. We build on Scotchmer (1985) and Pollak (1989) to develop the argument when consumers buy a basket of goods. The idea is simple. The fact that a consumer paid some premium to purchase a basket of goods implies that the consumer must have been willing to pay at least as much as that premium.

We combine ideas from the hedonic pricing literature (Nesheim, 2008; Bajari and Benkard, 2005) with revealed preference ideas

* Correspondence to: IFS, 7 RIdgmount Street, London WC1E 7AE, United Kingdom. Tel.: +44 2076795826.
E-mail addresses: rgriffith@ifs.org.uk (R. Griffith), l.nesheim@ucl.ac.uk (L. Nesheim).

0165-1765
http://dx.doi.org/10.1016/j.econlet.2013.04.040

from Blow et al. (2008) to analyse willingness to pay when consumers purchase continuous quantities of a high dimensional basket of goods. A major benefit of our approach is that we can exploit rich data without introducing aggregation bias and without making unnecessary separability assumptions. Under very minimal assumptions we are able to estimate bounds on willingness to pay; with more restrictive assumptions (but ones that are common in the literature) we can obtain point estimates of households' willingness to pay.

We illustrate our approach by estimating bounds on willingness to pay for organic foods using data on the shopping baskets of a large number of households. These estimates can inform regulation over the licencing 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.

2. Theoretical background

To develop intuition, we first describe bounds on willingness to pay in the single product case. Then we extend the analysis to the choice of a basket of products.

2.1. Demand for a single product

Let $z \in Z \subseteq \mathbf{R}_n$ be the vector of *all* product characteristics that affect consumer choice. Let $z(1)$, the first coordinate of z, be the 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 the equilibrium hedonic price.

Consider a consumer with characteristics x_h who buys a single unit of an organic product with product characteristics z^o and price p^o and elects not to buy a non-organic product with characteristics z^n and price p^n. Assume that the two products are identical in all dimensions other than organic. Let the consumer's indirect utility function be $v(x_h, z, p)$, where v is increasing in $z(1)$, continuously differentiable in p and strictly decreasing in p. If the consumer chooses the organic product, then revealed preference dictates

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

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

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

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

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

The left side of this expression is the willingness to pay for the organic characteristic. The right side is the organic price premium. For all consumers who buy organic, the price premium defines a lower bound on the willingness to pay for organic. For all consumers who do not buy organic, the price premium provides an upper bound on the willingness to pay for organic.

2.2. Demand for a basket

Let B_g be the set of products in category g and let $B = \bigcup_g B_g$ be the set of all products. For each product $b \in B_g$, let $z_b \in Z_g$ be its vector of characteristics. Define $\overline{z} = \{z_b\}_{b \in B}$ as the vector of characteristics of all products.

Let $\overline{v} = v\left(x_h, \overline{z}, \overline{p}\right)$ 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, $\overline{q}$, to minimise costs of attaining the fixed utility level $\overline{v}$. The consumer's total expenditure is

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

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

Denote what the consumer would have paid to obtain the same utility level if all products were converted to non-organic products with non-organic prices,

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

where $\overline{p}^n = \left\{p_b^n\right\}_{b \in B}$ and $\overline{z}^n = \left\{z_b^n\right\}_{b \in B}$ are the vectors of prices and characteristics in the counterfactual world where all products are converted to non-organic varieties. For household x_h, the willingness to pay for organic is the difference between these expenditures,

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

It is the negative of compensating variation.

If we assume that the utility function is known, then we can calculate a point estimate of willingness to pay using the price premia. More generally, if the utility function is not known, we cannot calculate willingness to pay. Nevertheless, 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

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

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, let $\overline{z}^o = \left\{z_b^o\right\}_{b \in B}$ 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 $p_b^o = h_g\left(z_b^o\right)$ 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,

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

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

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

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

2.3. Data

We use data from the 2004 Kantar Worldpanel for the UK to estimate (4) and (5). Households record purchases of all food, toiletries and cleaning products 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