Valuing a New Good

Laura Blow
Ian Crawford

THE INSTITUTE FOR FISCAL STUDIES
WP04/03

VALUING A NEW GOOD

Laura Blow and Ian Crawford
Institute for Fiscal Studies[1]

Abstract: This paper presents a nonparametric method for calculating a lower bound on the virtual or reservation price of a new good. This allows the welfare effects of product market innovations to be investigated. We illustrate the technique using consumer panel data.

Key Words: New goods, Revealed Preference
JEL Classification: C43, D11.

Acknowledgments: We are very grateful to Richard Blundell, Martin Browning, Peter Davis, J. Peter Neary, two anonymous referees and seminar participants at the NBER Productivity Meeting, University College Dublin, University College London and the Bank of England for their comments. This study was jointly funded by the Leverhulme Trust (Grant Ref: F/386/H), ESRC Centre for the Microeconomic Analysis of Public Policy at IFS and the AKF research project "Consumer Demand for Organic Foods - Domestic and Foreign Market Perspectives" financed by the Danish Ministry of Food, Agriculture and Fisheries. All errors are the sole responsibility of the authors

1. Introduction

The measurement of the welfare effects of the introduction of a new product in consumer markets is important in many areas in economics, from the correct measurement of cost-of-living indices to the evaluation of mergers and price regulation issues in industrial organisation[2]. The arrival of a new good is potentially welfare-improving because it expands the set of choices available to the consumer. Valuing this welfare change requires the identification of the reservation or 'virtual' price[3] of the good. The most common approach to calculating this is the parametric[4] estimation of a demand system, which is then extrapolated to solve for the virtual price[5]. Such methods have typically resulted in estimates of the welfare gains that are very large, which in turn has lead to worries that this may be due to functional form assumptions and the extrapolation of the demand curve outside of the range of any price variation that may have been observed. This paper presents an alternative nonparametric approach which uses revealed preference restrictions. At its heart, this method only requires the existence of a well-behaved

[1] Address for correspondence: Institute for Fiscal Studies, 7 Ridgmount Street, London, WC1E 7AE. laura_blow@ifs.org.uk or ian_crawford@ifs.org.uk

[2] See for example Boskin *et al* (1996) and the papers collected in T. F. Bresnahan and R. J. Gordon (ed's) (1997)

[3] The term is due to Rothbarth (1941).

[4] Heckman (1974) discusses the econometrics of virtual prices in a labour supply setting.

[5] Recent examples are Hausman (1997a,b) and Nevo (2001).

utility function. The form of the utility function need not be specified and no further restrictions are necessary. It can also make use of very few observations and does not seek to extrapolate from these. Finally, it is computationally very simple. We illustrate the idea with an application to a consumer panel dataset on milk.

2. New goods and consumer welfare

A new good[6] is usually thought of as a special case of a rationed good: non-existence is treated like a ration level of zero. Hicks (1940) and Rothbarth (1941) and more recently Neary and Roberts (1980) discuss the question of how to deal with rationed goods in economic problems. They show how the properties of demands under rationing can be expressed in terms of unrationed choices by replacing the observed market prices with a vector of 'virtual' prices or 'support' prices. Convexity, continuity and strict monotonicity of the consumer's preferences are sufficient to ensure that there always exists a set of strictly positive support prices consistent with any set of demands[7]. The virtual or support prices for the unrationed goods are identical to their actual prices[8] and so the term 'virtual' is therefore usually reserved for the support prices of the rationed goods only.

To place the new goods problem in a simple rationing context we suppose that there are $T+1$ periods, $t = 0, ..., T$, and $K+1$ goods, $k = 0, ..., K$. The 0th good is subject to a ration level of 0 in period $t = 0$ but is freely available from period 1 onwards. All other goods are freely available in every period. We denote by $\mathbf{q}_t^K$ and $\mathbf{p}_t^K$ the $(K \times 1)$ sub-vectors consisting of quantities and prices of the $k = 1, .., K$ goods in period t. The consumer's problem in period 0 is

$$\max_{\mathbf{q}} u(\mathbf{q}) \quad \text{subject to} \quad \mathbf{p}_0^{K\prime}\mathbf{q}^K \leq x_0 \ \text{ and } \ q_0^0 = 0 \tag{2.1}$$

and

$$\max_{\mathbf{q}} u(\mathbf{q}) \quad \text{subject to} \quad \mathbf{p}_t'\mathbf{q} \leq x_t \tag{2.2}$$

for all $t > 0$ where x_t denotes the available budget in period t. The first order conditions in the pre-introduction period are

$$\begin{bmatrix} u'\left(q_0^0\right) \\ \\ u'\left(\mathbf{q}_0^K\right) \end{bmatrix} = \lambda_0 \begin{bmatrix} \frac{\mu_0}{\lambda_0} \\ \\ \mathbf{p}_0^K \end{bmatrix} \tag{2.3}$$

and

$$u'\left(\mathbf{q}_t\right) = \lambda_t \mathbf{p}_t \tag{2.4}$$

otherwise. The scalar λ_0 is the marginal utility of income and μ_0 is the value of the rationing constraint in period 0. The vectors $\boldsymbol{\pi}_0 = \left[\frac{\mu_0}{\lambda_0}, \mathbf{p}_0^{K\prime}\right]'$ and $\boldsymbol{\pi}_t = \mathbf{p}_t$ for $t \neq 0$ are the support price vectors where the virtual price of the new good is

$$\pi_0^0 = \frac{\mu_0}{\lambda_0} \tag{2.5}$$

[6] In what follows we concentrate on the launch of a single new good. It will be clear in due course that our methods can be extended to the simultaneous launch of two or more goods. This would require the existence of a data-consistent utility function with a more restrictive (weakly separable) structure.

[7] Neary and Roberts (1980).

[8] Neary and Roberts (1980), p.27-9.

and

$$\pi_0^0 = \frac{u'\left(q_0^0\right)}{\lambda_0}$$

so that the virtual price measures the welfare gain associated with a marginal relaxation of the rationing constraint, normalised by the marginal utility of income at that point. The support price vectors for all the other periods are simply the observed prices. Clearly the support prices are such that the outcome of the rationed model is identical to the outcome of the unrationed choice generated by

$$\max_{\mathbf{q}_t} u\left(\mathbf{q}_t\right) \text{ subject to } \boldsymbol{\pi}_t'\mathbf{q}_t \leq x_t \text{ for } t = 0, 1, ..., T \tag{2.6}$$

in every period.

The most usual approach to calculating the virtual price of a new good has been the parametric estimation of demand curves. Generally an integrable demand system is posited and estimated using data from the post-introduction period. This is then used to "backcast" to the pre-introduction period by extrapolating the demand curve for the new good back to the $q^0 = 0$ axis. The price consistent with this demand (i.e. π_0^0) can then be read off. Due to worries about the high virtual prices often recovered by this method and their reliance on functional form assumptions we suggest a conservative nonparametric alternative based on revealed preference restrictions.

3. A revealed preference approach

The first attraction of revealed preference conditions is that they apply to any well behaved utility function and, beyond this, they require no additional restrictions on the precise form of preferences underlying consumer demands. This property is set out in Afriat's Theorem[9] which shows that, if a consumer's observed choices, given the prices they face, satisfy the Generalised Axiom of Revealed Preference (GARP), then these choices are consistent with having been generated by the maximisation of a well behaved utility function subject to a linear budget constraint[10]. The second attraction of the revealed preference approach which we are proposing is that it is computationally very simple. Finally, as we show, it can make effective use of *very* few post-introduction price observations. The drawback is that we recover a bound, not a point estimate. Nevertheless this is a conservative, useful and robust piece of information and one which does not derive from functional assumptions. Following Varian (1982) we set out the following definitions of revealed preference conditions;

Definition 1. $\mathbf{q}_t$ is directly revealed preferred to $\mathbf{q}$, written $\mathbf{q}_t R^0 \mathbf{q}$, if $\boldsymbol{\pi}_t'\mathbf{q}_t \geq \boldsymbol{\pi}_t'\mathbf{q}$.
Definition 2. $\mathbf{q}_t$ is strictly directly revealed preferred to $\mathbf{q}$, written $\mathbf{q}_t P^0 \mathbf{q}$, if $\boldsymbol{\pi}_t'\mathbf{q}_t > \boldsymbol{\pi}_t'\mathbf{q}$.
Definition 3. $\mathbf{q}_t$ is revealed preferred to $\mathbf{q}$, written $\mathbf{q}_t R \mathbf{q}$, if $\boldsymbol{\pi}_t'\mathbf{q}_t \geq \boldsymbol{\pi}_t'\mathbf{q}_s$, $\boldsymbol{\pi}_s'\mathbf{q}_s \geq \boldsymbol{\pi}_s'\mathbf{q}_r$, ..., $\boldsymbol{\pi}_m'\mathbf{q}_m \geq \boldsymbol{\pi}_m'\mathbf{q}$, for some sequence of observations $(\mathbf{q}_t,\ \mathbf{q}_s, ..., \mathbf{q}_m)$. In this case, we say that the relation R is the transitive closure of the relation R^0.

[9] Afriat (1967), Diewert (1973), Varian (1982).
[10] That is, there exists some well-behaved utility function $u(\mathbf{q})$ such that $u(\mathbf{q}_t) \geq u(\mathbf{q})$ for any $\mathbf{q}$ such that $\mathbf{p}_t'\mathbf{q}_t \geq \mathbf{p}_t'\mathbf{q}$.

Definition 4. $\mathbf{q}_t$ is strictly revealed preferred to $\mathbf{q}$, written $\mathbf{q}_t P \mathbf{q}$, if there exist observations $\mathbf{q}_s$ and $\mathbf{q}_m$ such that $\mathbf{q}_t R \mathbf{q}_s,\ \mathbf{q}_s P^0 \mathbf{q}_m,\ \mathbf{q}_m R \mathbf{q}$.
Definition 5. Data can be said to satisfy GARP if $\mathbf{q}_t R \mathbf{q}_s \Rightarrow \boldsymbol{\pi}_s' \mathbf{q}_s \leq \boldsymbol{\pi}_s' \mathbf{q}_t$. Equivalently, the data satisfy GARP if $\mathbf{q}_t R \mathbf{q}_s$ implies not $\mathbf{q}_s P^0 \mathbf{q}_t$.

Our aim is to use the restrictions imposed by revealed preference theory to place a lower bound on the virtual price of the new good in period 0. Suppose that we have data on prices and demands in period 0, $\{\boldsymbol{\pi}_0; \mathbf{q}_0\}$, with the missing element π_0^0 and that we also have data on prices and demands after the introduction of the new good,$\{\boldsymbol{\pi}_t, \mathbf{q}_t\}_{t=1,\ldots,T}$ with no missing variables. The lower bound on the virtual price consistent with utility maximising behaviour and a stable, well-behaved utility function is

$$\min \left\{ \pi_0^0 : \{\boldsymbol{\pi}_0, \boldsymbol{\pi}_t; \mathbf{q}_0, \mathbf{q}_t\}_{t=1,\ldots,T} \text{ passes GARP} \right\} \tag{3.1}$$

Given post-introduction data $\{\boldsymbol{\pi}_t; \mathbf{q}_t\}_{t=1,\ldots,T}$ which pass GARP, this can be computed in the following way. Let the sub-set of all post-introduction observations which are revealed preferred to $\mathbf{q}_0$ be denoted by

$$\mathbf{Q} = \{\mathbf{q}_t : \mathbf{q}_t R \mathbf{q}_0,\ t > 0\} \tag{3.2}$$

Then GARP requires not $\mathbf{q}_0 P^0 \mathbf{q}_t$ (i.e. $\boldsymbol{\pi}_0' \mathbf{q}_0 \leq \boldsymbol{\pi}_0' \mathbf{q}_t$) $\forall \mathbf{q}_t \in \mathbf{Q}$ and hence the lower bound on the virtual price is

$$\overline{\pi}_0^0 = \max_t \left\{ \min \pi_0^0 : x_0 \leq \boldsymbol{\pi}_0' \mathbf{q}_t,\ \forall \mathbf{q}_t \in \mathbf{Q} \right\} \tag{3.3}$$

where[11] the solution to the minimisation problem

$$\min \pi_0^0 : x_0 \leq \boldsymbol{\pi}_0' \mathbf{q}_t \tag{3.4}$$

is

$$\frac{\boldsymbol{\pi}_0^{K\prime} \left(\mathbf{q}_0^K - \mathbf{q}_t^K \right)}{q_t^0} \quad \text{if } q_t^0 > 0; \qquad 0 \text{ otherwise} \tag{3.5}$$

The number $\overline{\pi}_0^0$ is the lowest value of π_0^0 which makes the choice $q_0^0 = 0$ consistent with the unrationed maximisation of a stable utility function, i.e. precisely the virtual price of the new good in period 0 that we wish to calculate[12]. Note that if the post introduction data $\{\boldsymbol{\pi}_t; \mathbf{q}_t\}_{t=1,\ldots,T}$ violates GARP then there can exist no virtual price such that $\{\boldsymbol{\pi}_0, \boldsymbol{\pi}_t; \mathbf{q}_0, \mathbf{q}_t\}_{t=1,\ldots,T}$ passes GARP. Note too that if there are no post-introduction observations which are revealed preferred to $\mathbf{q}_0$ and where the consumer buys the new good, then there are no restrictions and hence no bound. Assuming, for the moment, that some $\overline{\pi}_0^0$ can be found, then we have the following result which shows that if we can recover $\overline{\pi}_0^0$, then we know that it is indeed the minimum value that π_0^0 can take given the requirement that the consumer is rational.

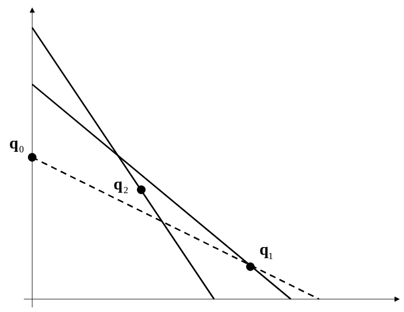
[11] Note that if $\mathbf{q}_t P^0 \mathbf{q}_0$ then the inequality is strict.
[12] Note that this bound could negative, in which case we would set it to zero, since we assume that consumption of the new good never detracts from utility, and hence that $\pi_0^0 \geqslant 0$.

Theorem 1: *If* $\underline{\pi}_0^0 < \overline{\pi}_0^0$ *then* $\{\boldsymbol{\pi}_0, \boldsymbol{\pi}_t; \mathbf{q}_0, \mathbf{q}_t\}_{t=1,\ldots,T}$ *violates GARP.*

Proof:

(1) Denote $\overline{\boldsymbol{\pi}}_0 = \left(\overline{\pi}_0^0, \pi_0^1, ..., \pi_0^K\right)$
(2) Let s index the observation which gives the maximum lower bound. Then $\overline{\pi}_0^0$ is such that $\overline{\boldsymbol{\pi}}_0'\mathbf{q}_s = \overline{\boldsymbol{\pi}}_0^{K\prime}\mathbf{q}_0^K = \overline{\boldsymbol{\pi}}_0'\mathbf{q}_0 = x_0$ where $\mathbf{q}_s R \mathbf{q}_0$
(3) Suppose we take $\underline{\pi}_0^0 < \overline{\pi}_0^0$, and let $\underline{\boldsymbol{\pi}}_0 = \left(\underline{\pi}_0^0, \pi_0^1, ..., \pi_0^K\right)$
(4) Then from (2) and (3) $\underline{\boldsymbol{\pi}}_0'\mathbf{q}_s < \overline{\boldsymbol{\pi}}_0'\mathbf{q}_s = \overline{\boldsymbol{\pi}}_0'\mathbf{q}_0 = \underline{\boldsymbol{\pi}}_0'\mathbf{q}_0$ (since $q_0^0 = 0$) $\Rightarrow \mathbf{q}_0 P^0 \mathbf{q}_s$ which is a violation of GARP.■

Figure 1

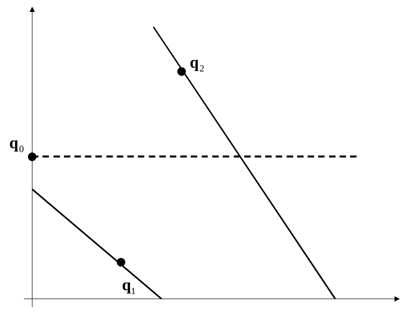


The basic idea is illustrated in Figure 1 in which there are three observations: the pre-introduction observation $\mathbf{q}_0$ (where the good measured on the horizontal axis has yet to be introduced), and two post-introduction observations $\mathbf{q}_1$ and $\mathbf{q}_2$. Our method essentially asks what are the limits that rationality places on the slope of the virtual budget line passing through $\mathbf{q}_0$? The answer is that any budget line through $\mathbf{q}_0$ steeper than the dashed line in Figure 1 is admissible. If we chose a lower virtual price for the new good then we would have $\mathbf{q}_0 P^0 \mathbf{q}_1$ and a violation of GARP. Thus we can nonparametrically bound the virtual price, using only the restriction of rationality and with very little data.

The main problem with these bounds is that they may not be informative if movements of the budget line between periods are large and relative price changes small. In these circumstances data may lack power either to reject[13], or usefully to invoke GARP. In particular any observations such that $\mathbf{q}_t \gg \mathbf{q}_0$ and any observations such that $\boldsymbol{\pi}_t'\mathbf{q}_t < \boldsymbol{\pi}_t'\mathbf{q}_0$ (including, in both cases, those observations where $q_t^0 = 0$) are uninformative. Such a situation is illustrated in Figure 2 in which the lower bound on the virtual price is zero.

[13]As pointed out by Varian (1982)

Figure 2



In these circumstances our method will not work well and we would need either more post-introduction data (in the hope that some observations will be informative) or more information from other sources. We now explore how additional information might help.

The key issue is that variation in the total budget between observations can remove the informational content of the data. If we were able to control the total budget at each observation what levels would we optimally set? We can show that, if demands are normal[14], we would ideally set the budgets at every post-introduction observation to be the lowest value which still yields a $\mathbf{q}_t$ that is revealed preferred to $\mathbf{q}_0$. Let $\mathbf{q}_t(x)$ denote the Marshallian demand vector at budget x, given period t prices. Then for each post-introduction observation we need to find the minimum x such that $\mathbf{q}_t(x)\, R\mathbf{q}_0$. The bound on the virtual price can then be set in a manner analogous to the original method, but using these demands. Let the set of all observations at these budget levels be denoted by

$$\widetilde{\mathbf{Q}} = \{\widetilde{\mathbf{q}}_t : \widetilde{\mathbf{q}}_t = \mathbf{q}_t(\min\{x : \mathbf{q}_t(x)\, R\mathbf{q}_0\})\} \tag{3.6}$$

Then the lower bound on the virtual price is

$$\widetilde{\pi}_0^0 = \max_t \left\{ \min \pi_0^0 : x_0 \leq \boldsymbol{\pi}_0' \widetilde{\mathbf{q}}_t,\ \forall \widetilde{\mathbf{q}}_t \in \widetilde{\mathbf{Q}} \right\} \tag{3.7}$$

where the solution to the minimisation problem

$$\min \pi_0^0 : x_0 \leq \boldsymbol{\pi}_0' \widetilde{\mathbf{q}}_t \tag{3.8}$$

is

$$\frac{\boldsymbol{\pi}_0^{K\prime}\left(\mathbf{q}_0^K - \widetilde{\mathbf{q}}_t^K\right)}{\widetilde{q}_t^0} \quad \text{if } \widetilde{q}_t^0 > 0; \qquad 0 \text{ otherwise.} \tag{3.9}$$

We then have the following result paralleling Theorem 1.

[14] That is, $\mathbf{q}(x)$ is a vector valued function such that $\mathbf{q}(x) \geq \mathbf{q}(x')$ for all $x \geq x'$.

Theorem 2: *If* $\pi_0^0 < \widetilde{\pi}_0^0$ *then* $\{\boldsymbol{\pi}_0, \boldsymbol{\pi}_t; \mathbf{q}_0, \widetilde{\mathbf{q}}_t\}_{t=1,\ldots,T}$ *violates GARP.*

Proof:

Analogous to Theorem 1. ■

We can also show that, if demands are normal, then demands at these particular budget levels will always provide a tighter bound on the virtual price than any other arbitrary set of demands.

Theorem 3: *If all goods are normal then* $\widetilde{\pi}_0^0 \geq \overline{\pi}_0^0$.

Proof:

(1) Suppose the bound $\overline{\pi}_0^0$ is provided by the observation $\mathbf{q}_s \in \mathbf{Q}$.

(2) From the definition of the set $\widetilde{\mathbf{Q}}$, $\boldsymbol{\pi}_s'\widetilde{\mathbf{q}}_s \leq \boldsymbol{\pi}_s'\mathbf{q}_s$ and therefore, by normality, $\widetilde{q}_s^k \leq q_s^k \ \forall k$.

(3) By definition $\overline{\pi}_0^0 = \boldsymbol{\pi}_0^{K\prime}\left(\mathbf{q}_0^K - \mathbf{q}_s^K\right)/q_s^0$, and, since $\widetilde{q}_s^k \leq q_s^k \ \forall k$, this implies $\overline{\pi}_0^0 = \boldsymbol{\pi}_0^{K\prime}\left(\mathbf{q}_0^K - \mathbf{q}_s^K\right)/q_s^0 \leq \boldsymbol{\pi}_0^{K\prime}\left(\mathbf{q}_0^K - \widetilde{\mathbf{q}}_s^K\right)/\widetilde{q}_s^0 \equiv \pi_0^0\left(\widetilde{\mathbf{q}}_s\right)$.

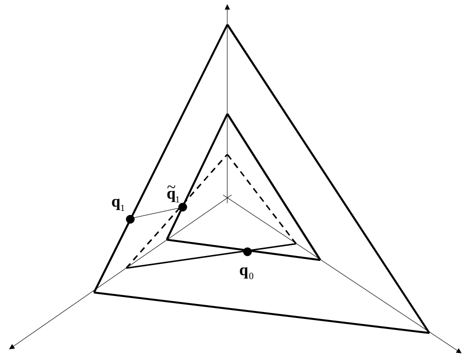
(4) By definition $\widetilde{\pi}_0^0 = \max_t \left\{\min \pi_0^0 : x_0 \leq \boldsymbol{\pi}_0'\widetilde{\mathbf{q}}_t, \ \forall \widetilde{\mathbf{q}}_t \in \widetilde{\mathbf{Q}}\right\}$, hence $\pi_0^0\left(\widetilde{\mathbf{q}}_s\right) \leq \widetilde{\pi}_0^0$.

(5) Thus $\overline{\pi}_0^0 \leq \pi_0^0\left(\widetilde{\mathbf{q}}_s\right) \leq \widetilde{\pi}_0^0$.■

If we are able to control the budget levels at each observation, then we can always potentially improve the bound (not least by increasing the number of informative observations by utilising observations not originally revealed preferred to $\mathbf{q}_0$). The question is how might we go about this? The simplest way is to make an assumption like local homotheticity. This is illustrated in Figure 3 where we have two observations $\{\mathbf{q}_0, \mathbf{q}_1\}$. The post-introduction observation $\mathbf{q}_1$ is not informative (the budget level is too high), but given we know the budget shares at prices $\boldsymbol{\pi}_1$ and the budget x_1, it is easy to find the informative demand $\widetilde{\mathbf{q}}_1$ by assuming that shares are roughly constant over the range $[\widetilde{x}_1, x_1]$. We can then derive $\widetilde{\pi}_0^0$ by noting that any virtual budget plane $\boldsymbol{\pi}_0'\mathbf{q} = x_0$ with its dashed sides steeper than those shown[15] would give a GARP violation for the data $\{\boldsymbol{\pi}_0, \boldsymbol{\pi}_1; \mathbf{q}_0, \widetilde{\mathbf{q}}_1\}$. It is not necessary, of course, to assume that demands are globally homothetic, and as long $\widetilde{x}_1$ and x_1 are not too far apart then local invariance in budget shares might be a reasonably inoffensive assumption to make. With more data (many repeat observations on the consumer's demands at different budgets with prices and other conditioning variables held constant) it will be possible to estimate the relationship nonparametrically by estimating a system of Engel curves.

[15] We have gone to a 3d example here because in 2d, or in 3d when there are no relative price movements, the planes $\boldsymbol{\pi}_0'\mathbf{q} = x_0$ and $\boldsymbol{\pi}_1'\mathbf{q} = \widetilde{x}_1$ coincide.

Figure 3



So far we have set aside the possibility that the post-introduction data violate GARP – other than by noting that if they do then our method, as described above, is not applicable. One option which would allow progress to be made in the face of GARP violations is based on weakening the requirements for rationality. A GARP test can be interpreted as a test of two sub-hypotheses[16]

1. the consumer has rational preferences.
2. the consumer is an efficient programmer.

If the data violates GARP then these hypotheses can be modified. Afriat's suggestion is that if (1) is not to be modified, then (2) must be modified. Instead of requiring exact efficiency, a form of partial efficiency is allowed, denoted by e where $0 \leq e \leq 1$. The consumer is now allowed to waste a fraction $(1 - e)$ of their budget through optimisation error. This is done by modifying the preference relation R^0 to

$$\mathbf{q}_s R_e^0 \mathbf{q}_t \Leftrightarrow e\boldsymbol{\pi}'_s \mathbf{q}_s \geq \boldsymbol{\pi}'_s \mathbf{q}_t \tag{3.10}$$

This stiffens the empirical requirement to reveal a preference in the sense that $\boldsymbol{\pi}'_s \mathbf{q}_s$ now has to be $1/e$ times bigger than $\boldsymbol{\pi}'_s \mathbf{q}_t$ in order to reveal a preference for $\mathbf{q}_s$ over $\mathbf{q}_t$. This efficiency concept can be used to define a weaker consistency test:

$$GARP\,(e) : \mathbf{q}_s R_e \mathbf{q}_t \Rightarrow \text{Not } \mathbf{q}_t P_e^0 \mathbf{q}_s \tag{3.11}$$

where "not $\mathbf{q}_t P_e^0 \mathbf{q}_s$" $\equiv e\boldsymbol{\pi}'_t \mathbf{q}_t \leq \boldsymbol{\pi}'_t \mathbf{q}_s$ and where R_e denotes the transitive closure of R_e^0. Note that if $e = 1$ then $GARP\,(e)$ is equivalent to $GARP$ and that if $e = 0$ then there is no restriction on behaviour. The ideas outlined above for bounding π_0^0 can be generalised to allow for inefficiency in the following way. The bound described in (3.3) becomes

$$\min \left\{ \pi_0^0 : \{\boldsymbol{\pi}_0, \boldsymbol{\pi}_t; \mathbf{q}_0, \mathbf{q}_t\}_{t=1,\dots,T} \text{ passes GARP}\,(e) \right\} \tag{3.12}$$

where e is the maximum value such that the post-introduction data $\{\boldsymbol{\pi}_t; \mathbf{q}_t\}_{t=1,\dots,T}$ pass GARP(e). We then denote the sub-set of observations revealed preferred to $\mathbf{q}_0$ at this e by

$$\mathbf{Q}_e = \{\mathbf{q}_t : \mathbf{q}_t R_e \mathbf{q}_0\} \tag{3.13}$$

[16] Afriat. S. N. (1973)

and the lower bound is

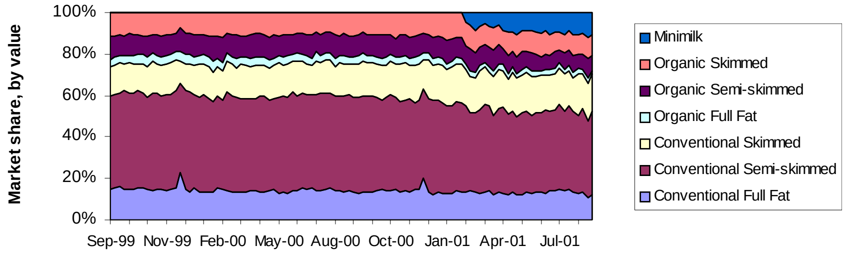
$$\overline{\pi}_0^0(e) = \max_t \left\{ \min \pi_0^0 : ex_0 \leq \boldsymbol{\pi}_0' \mathbf{q}_t, \ \forall \mathbf{q}_t \in \mathbf{Q}_e \right\} \tag{3.14}$$

which clearly reduces $\overline{\pi}_0^0$ as e moves below one. Whilst one might specify an acceptable level of e beforehand, it is also very simple to calculate the maximum value that e can take such that any given dataset satisfies GARP, and then to calculate the bound conditional on this level of efficiency. Our method of improving the bounds by controlling the budget levels at each observation also generalises to allow for cost inefficiency - a description of the algorithm we use can be found in the appendix.

4. An empirical illustration; "*Minimælk*"

A new type of milk was introduced in Denmark in the week beginning 5^{th}February 2001. This was organically produced and contained 0.5% fat and called *minimælk*/minimilk to distinguish it from skimmed milk (0.1% fat content) and semi-skimmed milk (1.5% fat). Figure 4 plots the market shares (by value) of the various milk types over the period September 1999 to September 2001. As can be seen minimilk quickly established a market share of about 10% by value.

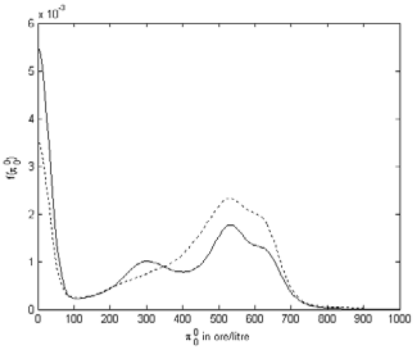
Figure 4: Percent market shares by value September 1999 to September 2001



Our dataset is an unbalanced 2 year Danish consumer panel of 1,614 milk-buying households observed over the period September 1999 to September 2001. The data records some demographic information and their weekly purchases of milk. In what follows we use data from the period after 5^{th} February 2001 to calculate the virtual price of organic minimilk in each period in which the household is observed prior to that date. For each household we first calculate the bound directly from their post-introduction observed demands, that is, we set $e^0 = \max(e : \{\boldsymbol{\pi}_0, \boldsymbol{\pi}_t; \mathbf{q}_0, \mathbf{q}_t\}_{t=1,\ldots,T}$ passes $GARP(e))$ and then set $\overline{\pi}_0^0 = \max_t\{\min \pi_0^0 : e^0 x_0 \leq \boldsymbol{\pi}_0' \mathbf{q}_t, \ \forall \mathbf{q}_t \in \mathbf{Q}_{e^0}\}$. We then also look at the improved bounds derived under the assumption of local homotheticity (using the method described in the appendix).

Figure 5: The density of the distribution of virtual prices minimilk

(a) All values (b) Positive values



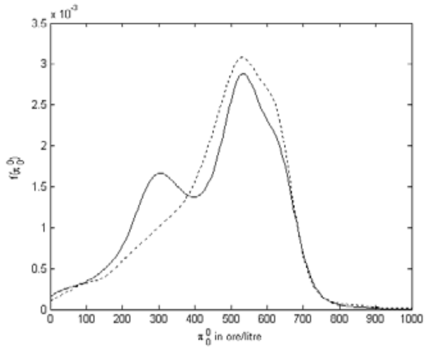


Figure 5 shows kernel estimates of the density of the distribution of virtual prices for minimilk[17]. The solid line shows the density recovered using our basic method, the dashed line shows the density when we tighten the bound. The right hand panel concentrates on households for which we are able to recover a non-zero bound. Further details are shown in Table 1. The improvement brought about by the local homotheticity assumption is apparent: there are fewer zero bounds, for example, in the improved set of bounds compared to the simple bounds (we can recover 74% non-zero bounds compared to 59% when the extra assumption is not made). In addition, the mean value of the improved bounds for the 17,867 observations for which we could get a positive simple bound is 495 øre/litre, as compared to 450 øre/litre from the simple bound. In calculating these bounds we allow for the minimum necessary level of optimisation error such that the post-introduction data pass GARP. We need to make very little such allowance: the median Afriat Efficiency required 0.999878.

Table 1: Descriptive statistics; bounds on the virtual price of minimilk

	Simple Bounds ($øre/litre$)		Improved Bounds ($øre/litre$)	
	All values	Positive values	All values	Positive values
Mean	263	450	354	476
Std Dev.	254.37	163.39	247.61	156.01
10th percentile	0	230	0	250
1st quartile	0	323	0	391
Median	265	495	436	504
3rd quartile	520	570	555	640
90th percentile	613	638	626	661
n	30504	17867	30504	22674

[17] The unit of observation underlying these pictures is a virtual price for a certain household at a certain date - there are therefore repeated virtual prices for individual households, and multiple observations on virtual prices at a point in time.

Table 2: Descriptive statistics: OLS regression of simple virtual price bounds on household characteristics

	Coefficient	*Standard Error*
Mean age of adults	7.302	*0.616*
(Mean age of adults)^2	-0.0667	*0.006*
# Children	-20.007	*1.782*
# Adults	-33.502	*2.391*
Female head of household	-11.586	*4.794*
An existing organic buyer	6.340	*4.202*
Pre-launch share of full fat	-110.874	*6.256*
Copenhagen	23.127	*2.872*
Zealand	6.531	*3.166*
Total milk budget in kroner	0.372	*0.093*
Constant	330.109	*14.987*

Table 2 gives a simple illustration of the how the virtual price bounds (we have used the straightforward unimproved bounds here) vary with some household characteristics observed in the dataset. The table shows the coefficients of an OLS regression of the virtual price on the listed characteristics and reports coefficients and standard errors (all of the explanatory variables are significant at 95%). The highest welfare gains seem to have been amongst middle aged households (the virtual price relationship is quadratic and peaks at about 55 years). The presence of increasing numbers of children in the household tends to decrease the welfare gain and does increasing numbers of adults (although at a slower rate). Households in which the head is female tend to have lower gains. If the household was a buyer of organic milk prior to the launch of minimilk (which is organic too) then their welfare gains tends to be higher, whereas if the household tended to buy a lot of full fat milk prior to the launch of minimilk (as measured by the share by quantity) their welfare gains are reduced (minimilk is a low fat variety). Finally there is a positive relationship with the household's total milk budget; households which spend more on milk overall tended to experience higher welfare gains associated with the launch of the new variety.

5. Conclusions

This paper presents a nonparametric method for calculating the lower bound on the virtual price of a new good. This bound is chosen such that the data are consistent with the Generalised Axiom of Revealed Preference and, therefore, it is also consistent with the maximisation of a well-behaved utility function. As a result this bound encompasses all parametric solutions which arise from fitting integrable demand systems to the same data. We also present a method for improving the bounds recoverable. We argue that this approach has three principal merits compared to parametric estimation. First, it does not require a maintained assumption regarding the form of the utility function. Second, it is computationally simple. Thirdly it can make efficient use of very few post-introduction price observations. We provide an empirical illustration using consumer panel data on milk purchases.

References

[1] Afriat, S.N. (1967), "The construction of utility functions from expenditure data", *International Economic Review*, **8**, 67-77.

[2] Afriat, S.N. (1973), "On a system of inequalities in demand analysis: an extension of the classical method", *International Economic Review*, **14**, 460-472.

[3] Blundell, R., Browning, M. and I. Crawford (2002), "Revealed preference and non-parametric Engel curves", *Econometrica,* **71**, 205-240.

[4] Boskin *et al* (1996), "Toward a more accurate measure of the cost of living", *Final Report to the Senate Finance Commission from the Advisory Commission to Study the Consumer Price Index.*

[5] Bresnahan, T. and R. J. Gordon (eds.) (1997), *The Economics of New Goods*, Chicago, University of Chicago Press.

[6] Diewert, W. E. (1973), "Afriat and Revealed Preference Theory", *Review of Economic Studies*, 40, 419-426

[7] Hausman, J. (1997a), "Valuation of new goods under perfect and imperfect competition", in Bresnahan, T. & Gordon, R.J., eds., *op cit.*

[8] Hausman, J. (1997b), "Celluar telephones, new goods and the CPI", *NBER* Working Paper Series, 5982.

[9] Heckman, J. (1974), "Shadow prices, market wages and labor supply", *Econometrica*, **42**, 679-694.

[10] Hicks, J. (1940), "The valuation of the social income", *Economica*, **7**, 105-124.

[11] Neary, J.P. and K.W.S. Roberts, (1980), "The theory of household behaviour under rationing", *European Economic Review*, **13**, 25-42.

[12] Nevo, A. (2001), "New products, quality changes and welfare measures computed from estimated demand systems", NBER Working Paper No. w8425, forthcoming in *The Review of Economics and Statistics*

[13] Rothbarth, E. (1941), "The measurement of changes in real income under conditions of rationing", *Review of Economic Studies*, **8**, 100-107.

[14] Varian, H. (1982), "The non-parametric approach to demand analysis", *Econometrica*, **50**, 945-973.

A. Improving the bound

During the period of our data it was only possible to buy milk in discrete jumps of 0.25 litres because of the way it was packaged and so, when making use of the local hometheticity assumption, we respect this and only use observations $\widehat{\mathbf{q}}_t$ defined from the observed set of demands $\{\mathbf{q}_t\}_{t=1,\dots,T}$ as follows:

$$\begin{aligned} \widehat{\mathbf{q}}_t &= m\mathbf{q}_t \text{ and } \widehat{\mathbf{q}}_t = 0.25\mathbf{i} \\ & \quad \text{where} \\ m &> 0 \text{ and } \mathbf{i} \text{ is a } (K \times 1) \text{ vector of zeros or positive integers} \end{aligned}$$

Call this set of demands $\widehat{\mathbf{Q}}$. Let $e^0 = \max_e \left\{ e : \{\boldsymbol{\pi}_t, \mathbf{q}_t\}_{t=1,\dots,T} \text{ satisfies } GARP\,(e) \right\}$. Dividing the original observations into two sets, $\mathbf{Q}_{e^0} = \{\mathbf{q}_t : \mathbf{q}_t R_{e^0} \mathbf{q}_0\}$ (as before) and $\mathbf{Q}^- = \{\mathbf{q}_t\}_{t=1,\dots,T} \setminus \mathbf{Q}_{e^0}$ (i.e. those observations revealed preferred to $\mathbf{q}_0$ with efficiency e^0, and those not revealed preferred to $\mathbf{q}_0$) we calculate the improved bound $\widehat{\pi}_0^0$ using the following algorithm:

Algorithm :

1. Set $\widehat{\pi}_0^0 = \max_t \left\{ \min \pi_0^0 : e^0 x_0 \leq \boldsymbol{\pi}_0' \mathbf{q}_t,\ \forall \mathbf{q}_t \in \mathbf{Q}_{e^0} \right\}$
2. Define the set $\mathbf{H}$ as $\max\limits_{\forall t : \mathbf{q}_t \in \mathbf{Q}_{e^0}} \left(\widehat{\mathbf{q}}_t \in \widehat{\mathbf{Q}} : m < 1, \widehat{\mathbf{q}}_t \notin \mathbf{Q}_{e^0} \right)$. Set $\mathbf{Q}_{e^0} = \mathbf{Q}_{e^0} \cup \mathbf{H}$.
3. Define the set $\mathbf{H}^-$ as $\min\limits_{\forall t : \mathbf{q}_t \in \mathbf{Q}^-} \left(\widehat{\mathbf{q}}_t \in \widehat{\mathbf{Q}} : m > 1, m \leq \min \left(m : e^0 \boldsymbol{\pi}_t' \widehat{\mathbf{q}}_t \geq \boldsymbol{\pi}_t' \mathbf{q}_0 \right), \widehat{\mathbf{q}}_t \notin \mathbf{Q}^- \right)$. Set $\mathbf{Q}^- = \mathbf{Q}^- \cup \mathbf{H}^-$. If $\mathbf{H}^- = \varnothing$ go to step 4, otherwise go to step 5.
4. Define the set $\mathbf{H}^-$ as $\max\limits_{\forall t : \mathbf{q}_t \in \mathbf{Q}^-} \left(\widehat{\mathbf{q}}_t \in \widehat{\mathbf{Q}} : m < 1, \widehat{\mathbf{q}}_t \notin \mathbf{Q}^- \right)$. Set $\mathbf{Q}^- = \mathbf{Q}^- \cup \mathbf{H}^-$.
5. If $\mathbf{H} = \varnothing$ and $\mathbf{H}^- = \varnothing$, stop.
6. Calculate $e^N = \max \left(e : \left\{ \boldsymbol{\pi}_0, \boldsymbol{\pi}_t; \mathbf{q}_0, \mathbf{Q}_{e^0}, \mathbf{Q}^- \right\}_{t=1,\dots,T} \text{ passes } GARP\,(e) \right)$. If $e^N < e^0$, stop.
7. Calculate the set $\mathbf{Q}^R = \{\mathbf{q}_t : \mathbf{q}_t R_{e^0} \mathbf{q}_0\ ,\ \forall \mathbf{q}_t \in \mathbf{Q}_{e^0} \cup \mathbf{Q}^-\}$
8. Set $\widehat{\pi}_0^0 = \max_t \left\{ \min \pi_0^0 : e^0 x_0 \leq \boldsymbol{\pi}_0' \mathbf{q}_t,\ \forall \mathbf{q}_t \in \mathbf{Q}^R \right\}$. Go to step 2.

In words, for each period for which the original observation was revealed preferred to $\mathbf{q}_0$ we move *in* along the expansion path adding successive feasible demands (step 2). For each period for which the original observation was not revealed preferred to $\mathbf{q}_0$ we move *out* along the expansion path adding successive feasible demands stopping as soon as we reach a demand that is directly revealed preferred to $\mathbf{q}_0$ (step 3). If this point is reached, we then start moving *in* along the expansion path from the original observed demand for completeness (step 4). After we add each new round of observations we check to see whether the e required for this updated dataset to pass GARP has decreased. If it has, then we stop, as our goal is to improve the bound on π_0^0 as far as possible without weakening the original efficiency, e^0, required for the data to pass GARP. If it has not then we use the updated data to calculate a new bound on π_0^0. It can be seen that the algorithm is finite as we can not go in along the expansion path further that the origin, or out further than the first demand that is directly revealed preferred to $\mathbf{q}_0$, and so we must achieve $\mathbf{H} = \varnothing$ and $\mathbf{H}^- = \varnothing$ in a finite number of iterations since there are only a finite number of feasible demands on any bounded section of the expansion path. We also have the following theorem:

Theorem 4: *If the data at all points on the expansion paths satisfy GARP perfectly then (given the discreteness of the possible demands allowed) the output $\widehat{\pi}_0^0$ from the algorithm must always equal the best possible bound $\widetilde{\pi}_0^0$ as defined in Section 3.*

Proof: Since $e^0 = e^N = 1$, the algorithm continues until $\mathbf{H} = \varnothing$ and $\mathbf{H}^- = \varnothing$. Therefore, when the algorithm stops, $\mathbf{Q}_{e^0}$ contains all the original observed $\mathbf{q}_t$ such that $\mathbf{q}_t R_{e^0} \mathbf{q}_0$ and all the feasible $0 < \widehat{\mathbf{q}}_t < \mathbf{q}_t$ along the expansion paths for these periods. For the periods where the original observations were not revealed preferred to $\mathbf{q}_0$, $\mathbf{Q}^-$ contains $\underline{\widehat{\mathbf{q}}}_t \equiv \min\left(\widehat{\mathbf{q}}_t : \widehat{\mathbf{q}}_t R_{e^0}^0 \mathbf{q}_0\right)$ and all the feasible $0 < \widehat{\mathbf{q}}_t < \underline{\widehat{\mathbf{q}}}_t$ for these periods. Hence, for each $t = 1, ..., T$, $\exists\ \mathbf{q}_t \in \mathbf{Q}_{e^0} \cup \mathbf{Q}^-$ such that $\mathbf{q}_t R_{e^0} \mathbf{q}_0$, and $\mathbf{Q}_{e^0} \cup \mathbf{Q}^-$ contains all feasible demands smaller than this $\mathbf{q}_t$. Hence, by the definition of $\widetilde{\mathbf{q}}_t$, $\mathbf{Q}_{e^0} \cup \mathbf{Q}^-$ must contain $\{\widetilde{\mathbf{q}}_t\}_{t=1,...,T}$.■