ISSN 1471-0498

DEPARTMENT OF ECONOMICS

DISCUSSION PAPER SERIES

GOODS VERSUS CHARACTERISTICS: REVEALED PREFERENCE PROCEDURES FOR NESTED MODELS

Matthew Polisson

Number 531
February 2011

Manor Road Building, Oxford OX1 3UQ

FEBRUARY 4, 2011

GOODS VERSUS CHARACTERISTICS: REVEALED PREFERENCE PROCEDURES FOR NESTED MODELS[*]

MATTHEW POLISSON[a]

This paper compares the goods and characteristics models of the consumer within a traditional demand framework. We examine the nonparametric revealed preference conditions for the goods and characteristics models, and we develop a methodology for testing nested models of this class using nonparametric revealed preference techniques. Of primary interest is to make a comparison on the basis of predictive success, which requires that we develop a method to relate set predictions across models. This allows us to nonparametrically identify the model which best fits the data, and in doing so, to identify the value added by the characteristics structure in explaining consumer behavior. We then explore the effects of hypothetical price variation as implied by our findings in order to nonparametrically bound any comparative statics of interest. We implement these procedures on household panel data from the UK milk market. The primary result is that the better fit of the characteristics model is entirely attributable to dimension reduction.

JEL CODES: D11, D12.

KEYWORDS: Characteristics, demand, dimension reduction, nested models, revealed preference.

[*]I am grateful to Martin Browning, Ian Crawford, Rachel Griffith, Valérie Lechene, John Quah, and seminar participants at the IFS and Oxford for helpful comments. I am also grateful to Kantar and the IFS for providing access to the data. Funding from the ESRC/MRC/NIHR, Reference Number G0802297, is gratefully acknowledged. All errors are my own.

[a]Department of Economics, University of Oxford, Manor Road, Oxford, OX1 3UQ and Institute for Fiscal Studies, 7 Ridgmount Street, London, WC1E 7AE

matthew.polisson@economics.ox.ac.uk

1. INTRODUCTION

This paper compares two models that are central to empirical work in applied microeconomics, namely the goods model of the consumer (see, for example, Deaton and Muellbauer (1980) and Banks, Blundell, and Lewbel (1997)) and the characteristics model of the consumer (see, for example, McFadden (1973) and Berry, Levinsohn, and Pakes (1995)). The former posits that an individual has preferences over goods, and the latter that an individual has preferences over characteristics, giving rise to derived preferences over goods (see, for example, Gorman (1980) and Lancaster (1966) for early descriptions of the characteristics model). Both models are parsimonious and empirically falsifiable, but the characteristics model has the further advantage of being able to reduce the dimensions of the consumer choice problem.

However, since the characteristics model is a special case of the goods model, it necessarily performs weakly worse from a *consistency* standpoint. It is well known that within a traditional demand framework, additional theoretical structure begets additional empirical restrictions, which then implies that a general model is at least as likely as a special case to be consistent with the data. While this raises a broader question about when to add structure to a model more generally, in this paper we focus exclusively on the goods and characteristics models. We develop a methodology for testing between them by assessing consistency in light of the criteria for consistency, and then by making a comparison across models. In doing so, we hope to identify the value added by the characteristics structure in explaining consumer behavior.

The conventional econometric approach has been to first impose functional form and heterogeneity restrictions, and then to estimate both models, comparing fit. However, even with flexible functional forms and allowing for heterogeneity, this approach necessarily conflates primitive and ancillary hypotheses. Instead, we proceed empirically in the tradition of nonparametric revealed preference, for which there is a well-established methodology of falsification and a vast supporting literature (see, for example, Afriat (1967), Diewert (1973), and Varian (1982)). Necessary and sufficient conditions are well-defined and easy to implement for both the goods and characteristics models (see Blow, Browning, and Crawford (2008) for the latter). These techniques exhaust the pure empirical implications of each theory and avoid

the identification problems associated with conventional parametric approaches.

A drawback of nonparametric revealed preference procedures is that they very often fail to give rise to stringent, or even non-trivial, empirical restrictions. Without sufficient price relative to expenditure variation, the ability to falsify is minimal. In a non-statistical sense, nonparametric revealed preference conditions very often lack restrictive (and predictive) *power*. This problem is well-known, and several treatments have been advanced, most notably by Bronars (1987). However, in this paper we are able to exploit the power differential between the goods model and the weakly more restrictive characteristics model. We make a comparison on the basis of predictive success, which requires that we develop a method to relate set predictions across models. To do this, we make use of an axiomatic approach developed by Selten (1991) and already extended to the revealed preference framework by Beatty and Crawford (2009). We relate consistency and power in order to establish *fit*, again in a non-statistical sense, and then make a comparison. Once we have identified the model which best fits the data, we explore the effects of hypothetical price variation as implied by our findings in order to nonparametrically bound any comparative statics of interest.

We implement these procedures on household panel data from the UK milk market. This highlights a relevant nutritional application, in which we aim to identify the relative importance of nutrients in explaining food consumption. There are an abundance of alternative markets, e.g., housing, automobile, education, and labor, but we focus on milk to simplify the implementation and to establish continuity with previous work in this area. Our data are from the Kantar UK Worldpanel, which is comprised of scanner data from a representative rolling household panel. For the sake of simplicity, we focus our analysis on the three main types of conventional non-organic milk—whole, semi-skimmed, and skimmed—and their defining nutritional attributes—"milkiness" and fat.

Our results suggest that on average the characteristics model outperforms the goods model in terms of fit, which then allows for a tightening of the bounds on own- and cross-price demand responses in some cases. However, the better fit of the characteristics model is entirely attributable to its dimension-reducing properties, which suggests that other dimension-reducing structures, e.g., weakly separable or

alternative characteristics structures, may have performed at least as well. Although there is a great deal of heterogeneity in our results, we find that consistency, power, and fit are largely unexplained by observed household attributes.

The contributions of this paper are threefold. Firstly, we are able to compare the goods and characteristics models completely nonparametrically on the basis of fit, which relates consistency and power across models. To our knowledge, this is the first time this has been achieved. Secondly, in order to make this comparison, we extend a power concept already developed in the revealed preference literature to the characteristics model, which does not follow immediately or trivially from previous work. Thirdly, we introduce a method that controls for dimension reduction when making model comparisons in general, which is critical in light of the dimension-reducing properties of the characteristics structure. Lastly, we emphasize that the methods developed here can, in principle, be applied to models of a similar class more generally, e.g., weakly separable, latently separable, and collective models.

The organization of this paper is as follows. Section 2 develops the basic theory. We first situate the analysis within a traditional demand framework and describe the critical features and empirical implications of the goods and characteristics models. We then outline the revealed preference conditions for these models, as well as a power concept already developed for the goods model. Next, we extend this concept to the characteristics model, and in doing so, we discuss our priors and how to handle dimension reduction. We then briefly review the axiomatization of predictive success. Lastly, we introduce hypothetical price variation in order to recover bounds on demand responses. Section 3 outlines the empirical implementation. Here we introduce the nutritional application and describe the data. Section 4 presents the results. Section 5 concludes.

2. THEORY

2.1. *Framework*

We begin by comparing the goods and characteristics models of the consumer. The primal problem in the goods model is to

$$\max_q u(q) \text{ subject to } p'q \leq x, \tag{2.1}$$

where $q \in \mathcal{R}_+^K$ are consumption goods, $p \in \mathcal{R}_{++}^K$ are prices, and $x \in \mathcal{R}_{++}$ is income, and where the utility function $u(\cdot) : \mathcal{R}_+^K \to \mathcal{R}$ is differentiable, non-satiated, and concave in q. The primal problem in the characteristics model is to

$$\max_z v(z) \text{ subject to } z = F(q) \text{ and } p'q \leq x, \tag{2.2}$$

where $z \in \mathcal{R}_+^J$ are characteristics, and where the utility function $v(\cdot) : \mathcal{R}_+^J \to \mathcal{R}$ is differentiable, non-satiated, and concave in z, and the mapping from goods to characteristics $F(\cdot) : \mathcal{R}_+^K \to \mathcal{R}_+^J$ is differentiable, nondecreasing, and concave in q. Convention suggests a linear mapping $F(q) = A'q$, where $A \in \mathcal{R}_+^{K \times J}$. We adopt a linear form but note that what follows applies to more general technologies. We further note that $J < K$ and that A has full rank by construction.

The solutions to (2.1) and (2.2) are the demand systems $q(p, x)$ and $\tilde{q}(p, x, A)$, respectively, where $q(\cdot) : \mathcal{R}_{++}^K \times \mathcal{R}_{++} \to \mathcal{R}_+^K$ and $\tilde{q}(\cdot) : \mathcal{R}_{++}^K \times \mathcal{R}_{++} \times \mathcal{R}_+^{K \times J} \to \mathcal{R}_+^K$ are both differentiable and zero homogenous in (p, x), satisfy adding-up, and give rise to a Slutsky matrix that is symmetric and negative semi-definite. However, there is a further empirical restriction that the number of goods consumed cannot exceed the number of characteristics, unless characteristics budget sets are uniquely aligned, which is neatly encapsulated by the rank condition

$$rank(A^+ \sim p^+) = rank(A^+), \tag{2.3}$$

where p^+ is a sub-vector of p and A^+ is a sub-matrix of A for which the corresponding $q_k > 0$, and where $\sim$ denotes the horizontal concatenation of A^+ and p^+. Since $J < K$ by construction, the rank condition normally restricts consumption to fewer than the full set of goods (see Appendix A.1 for further details on the empirical implications of the characteristics model).

Preferences over characteristics, combinations of which are embodied by goods, imply derived preferences over goods, i.e., there exists a derived utility function $\tilde{u}(\cdot) : \mathcal{R}_+^K \to \mathcal{R}$ such that $\tilde{u}(q) = v(A'q)$, which means that the characteristics model is a special case of the goods model. In this way, the goods model nests the characteristics model, and as consequence, any observable empirical restrictions are also nested.[1] We exploit this nested structure to identify which model better fits the data.

[1]We focus on characteristics models because they are widely used in applied empirical work (see, for example, McFadden (1973) and Berry, Levinsohn, and Pakes (1995)), and because they

2.2. *Revealed Preference Conditions*[2]

Since we observe prices and quantities, and hence expenditures, the restrictions outlined in the previous section are, in principle, empirically falsifiable. The conventional econometric approach has been to first impose functional form and heterogeneity restrictions. However, parametric techniques necessarily conflate primitive and ancillary hypotheses. Instead, we proceed empirically in a nonparametric revealed preference sense, for which there is a well-established methodology of falsification and a vast supporting literature (see, for example, Samuelson (1938, 1948), Houthakker (1950), and Richter (1966) for original formulations of the argument, Afriat (1967) and Diewert (1973) for necessary and sufficient conditions under which a set of observations is consistent with the standard goods model of the consumer, and Varian (1982, 1983) for operationalizable and computationally feasible tests).

2.2.1. *Goods Model*

Let $q_t \in \mathcal{R}^K_+$ be observed goods bought at prices $p_t \in \mathcal{R}^K_{++}$ in period $t \in \{1, \ldots, T\}$. If the set of observations (p_t, q_t), $t = 1, \ldots, T$, is consistent[3] with the goods model of the consumer, it can be rationalized as follows:

DEFINITION 2.1 A utility function $u(\cdot) : \mathcal{R}^K_+ \to \mathcal{R}$ *q*-rationalizes the set of observations (p_t, q_t), $t = 1, \ldots, T$, if $u(q_t) \geq u(q)$ for all q such that $p_t' q_t \geq p_t' q$.

The necessary and sufficient conditions for q-rationalization and the Generalized Axiom of Revealed Preference (GARP) are related by Afriat's Theorem (see Afriat (1967)) as follows:

THEOREM 2.1 *The following statements are equivalent:*

1. *There exists a utility function $u(\cdot) : \mathcal{R}^K_+ \to \mathcal{R}$ that is continuous, increasing, and concave which q-rationalizes the set of observations (p_t, q_t), $t = 1, \ldots, T$.*

are straightforward to implement nonparametrically (see Blow, Browning, and Crawford (2008)). However, nested restrictions enable us to identify the relative importance of additional structure more generally (see Appendix A.2 for further discussion on the empirical implications of nested models).

[2]See Appendix A.3 for revealed preference definitions and axioms, full versions of the theorems, and some brief discussion.

[3]By "consistent" we mean "does not falsify."

2. *There exist* $U_t \in \mathcal{R}$ *and* $\lambda_t \in \mathcal{R}_{++}$, $t = 1, \ldots, T$, *such that*

$$U_s \leq U_t + \lambda_t p_t'(q_s - q_t) \ \forall \ s, t = 1, \ldots, T.$$

3. *The set of observations* (p_t, q_t), $t = 1, \ldots, T$, *satisfies GARP.*

PROOF: See Afriat (1967), Varian (1982, 1983), and Fostel, Scarf, and Todd (2004). *Q.E.D.*

2.2.2. *Characteristics Model*

Let $z_t \in \mathcal{R}_+^J$ be observed characteristics bought at prices $p_t \in \mathcal{R}_{++}^K$ in period $t \in \{1, \ldots, T\}$, and let $A \in \mathcal{R}_+^{K \times J}$ be a linear technology mapping from q_t to z_t according to $z_t = A'q_t \ \forall \ t = 1, \ldots, T$. If the set of observations (p_t, q_t), $t = 1, \ldots, T$, and the technology A are consistent with the linear characteristics model of the consumer, they can be rationalized as follows:

DEFINITION 2.2 A utility function $v(\cdot) : \mathcal{R}_+^J \to \mathcal{R}$ z-rationalizes the set of observations (p_t, q_t), $t = 1, \ldots, T$, for the technology A if $v(z_t) = v(A'q_t) \geq v(A'q) = v(z)$ for all q such that $p_t'q_t \geq p_t'q$.

The necessary and sufficient conditions for z-rationalization are related by Blow, Browning, and Crawford (2008) as follows:

THEOREM 2.2 *The following statements are equivalent:*

1. *There exists a utility function* $v(\cdot) : \mathcal{R}_+^J \to \mathcal{R}$ *that is continuous, non-satiated, and concave which* z*-rationalizes the set of observations* (p_t, q_t), $t = 1, \ldots, T$, *for the technology* A.
2. *There exist* $U_t \in \mathcal{R}$, $\rho_t \in [1, \infty)$, *and* $\sigma_t \in \mathcal{R}^J$, $t = 1, \ldots, T$, *such that*

$$U_s \leq U_t + \sigma_t'(A'q_s - A'q_t) \ \textit{and}$$
$$\rho_t p_t \geq A\sigma_t \ \forall \ s, t = 1, \ldots, T,$$

the latter binding for any $q_{kt} > 0$.

PROOF: See Blow, Browning, and Crawford (2008). *Q.E.D.*

So far we have assumed that all prices are observed. In empirical practice, it is often the case that we only observe the prices of goods purchased, and then we impute the prices of goods not purchased. Blow, Browning, and Crawford (2008) develop a set of revealed preference conditions for the characteristics model that allow for missing prices by removing the inequality restrictions on the prices of goods not purchased in statement (2) of Theorem 2.2. The conditions are necessarily weaker since the prices of goods not purchased can be set arbitrarily high. Blow, Browning, and Crawford (2008) further note that it is possible to construct virtual prices, where individuals are just on the verge of buying, and also that setting the technology A equal to the identity matrix gives rise to the revealed preference conditions for the goods model with missing prices. In what follows, we assume that all prices are observed unless otherwise explicitly noted.

2.3. *Area Power*

Since the goods model nests the characteristics model, z-rationalization implies q-rationalization.[4] As a consequence, when we aggregate over a heterogeneous sample in the implementation, the pass rate for the goods model is always at least as high as the pass rate for the characteristics model. In other words, the characteristics model necessarily performs weakly worse than the goods model from a consistency standpoint. The central challenge of this paper is to formally identify the value added by the characteristics structure in light of these pass rates. In order to do so, we explore power differentials in the revealed preference conditions for the goods and characteristics models.

The necessary and sufficient revealed preference conditions outlined in the previous section are exhaustive, meaning that no further derivable empirical implications exist unless we impose additional structure. In choosing not to do so, we avoid conflating hypotheses, and we allow for maximal heterogeneity, both of which are desirable. However, the power of these tests depends vitally upon variations in the data. More specifically, we require sufficient price relative to expenditure variation such that

[4]More specifically, z-rationalization with imputed/missing prices implies q-rationalization with imputed/missing prices, respectively. Furthermore, z-rationalization with imputed prices implies z-rationalization with missing prices, and q-rationalization with imputed prices implies q-rationalization with missing prices.

implied budget hyperplanes intersect, which has the potential to generate cycles, and hence the power to falsify. This problem has long been noted, and several treatments have been advanced (see, for example, Andreoni and Harbaugh (2008) for a review). Nonetheless, we are able exploit the structure/power differential between the goods and characteristics models in order to establish comparative fit. The method we adopt is similar in spirit to that of Bronars (1987), but very closely follows upon Beatty and Crawford (2009), who have recently developed an axiomatic approach based on Selten (1991).

To illustrate, we begin with a simple two-good, two-period numerical example. Let $q_1 = (0, 14)'$ at $p_1 = (7, 6)'$ and $q_2 = (14, 0)'$ at $p_2 = (6, 7)'$. As shown in Figure 1, the set of observations is q-rationalizable, and since the implied budget lines intersect, this is non-trivial.

[Figure 1 about here]

In fact, at the same prices and implied expenditures, we can define a set of outcomes that are not q-rationalizable according to

$$(2.4) \qquad (q_1, q_2) : q_{11} \in (84/13, 12],\ q_{12} \in [0, 84/13),\ (7, 6)'q_1 = 84,\ (6, 7)'q_2 = 84.$$

Translating into budget shares gives

$$(2.5) \qquad (w_{11}, w_{12}) : w_{11} \in (7/13, 1],\ w_{12} \in [0, 6/13),$$

where $w_{11}, w_{12} \in [0, 1]$ are budget shares for good 1 in periods 1 and 2, respectively. The benefit of using budget shares is that with only two goods and two periods, the set of q-rationalizable outcomes, implied by adding-up, can be illustrated diagramatically as in Figure 2.

[Figure 2 about here]

Following Beatty and Crawford (2009), we define a relative *area* $a \in [0, 1]$ as the size of the set of q-rationalizable outcomes relative to the size of the set of all feasible outcomes. Loosely speaking, an area is a theory-consistent acceptance region, or target, that lies within the feasible outcome space. In this example, $a = 133/169 \approx 0.787$, which means that about 21.3% of feasible outcomes are ruled out by the revealed preference restrictions.

It is worth being precise about what area power measures. In the data, we only observe prices and quantities, which are all that we require to establish q-rationalizability. However, in order to calculate an area, we must first define a set of feasible outcomes. At the observed prices in a given period, we can define a set of goods bundles just as affordable as the observed goods bundle. In other words, the set of feasible goods bundles in a given period includes all points on the budget hyperplane implied by observed prices and quantities in that period. The feasible outcome space is simply the cartesian product of these sets across periods, i.e.,

$$\mathcal{F} = \times_{t=1}^{T} \mathcal{F}_t = \times_{t=1}^{T} \{q : p_t' q = p_t' q_t,\ q \in \mathcal{R}_+^K\}, \tag{2.6}$$

where $\mathcal{F}$ denotes the feasible outcome space and $\mathcal{F}_t$ denotes the set of feasible goods bundles in period t. Beatty and Crawford (2009) therefore note that area power is vitally dependent upon any priors about the feasible outcome space. For example, the set of feasible outcomes as just described may contain outcomes that are behaviorally irrelevant, and we may want to truncate the feasible outcome space accordingly. The area power treatment is therefore not so easily applied to the characteristics model, and indeed to many other models of a similar class, each of which require careful consideration of the feasible outcome space. We discuss our priors at length in the following section.

Lastly, note that in empirical practice we use Monte Carlo simulations to numerically estimate areas. In the two-good, two-period example above, we are able to construct an area geometrically and solve for it analytically. We use this simple example to develop intuition. However, with additional goods and periods, or with additional complicating structure, we use numerical methods. An analogous numerical procedure draws outcomes uniformly from implied budget hyperplanes and tests for q-rationalization until convergence is reached. While we require a probabilistic assumption to implement this procedure, what we are trying to estimate is the relative size of a *set* of theory-consistent outcomes. As Beatty and Crawford (2009) explicitly note, area power is not, in concept, a probability measure, though it has the requisite properties. When given this interpretation, it resembles statistical power (see Bronars (1987)),[5] which has been criticized for considering, as the hypothesis

[5]In fact, the area power measure is equal to one minus the statistical power measure of Bronars (1987).

alternative to the goods model of the consumer, that all outcomes on the implied budget hyperplanes are events occurring with equal probabilities. Such criticism is a reminder that area power critically depends upon the set of feasible outcomes against which the model of interest is being compared.

2.4. *Feasible Outcomes*

Here we discuss our priors over the set of feasible outcomes for the goods and characteristics models. First consider

(P1) $\mathcal{F} = \times_{t=1}^{T} \mathcal{F}_t = \times_{t=1}^{T} \{q : p_t' q = p_t' q_t,\ q \in \mathcal{R}_+^K\}$.

This corresponds to the set of feasible outcomes outlined in the previous section. Consider a simple three-good, two-characteristic, two-period numerical example. Let $q_1 = (0, 0, 14)'$ at $p_1 = (7, 4, 6)'$ and $q_2 = (14, 0, 0)'$ at $p_2 = (6, 4, 7)'$. The technology mapping from goods to characteristics is represented by

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 1 \\ 2 & 1 \end{pmatrix}. \tag{2.7}$$

The set of observations is q-rationalizable, and since the implied budget planes intersect, this is again non-trivial. Under (P1), $a^q \approx 0.816$, where a^q denotes the area power of the goods model, which means that about 18.4% of feasible outcomes are ruled out by corresponding revealed preference restrictions. Furthermore, as shown in Figure 3, the set of observations also non-trivially satisfies the revealed preference conditions for the characteristics model.

[Figure 3 about here]

Under (P1), $a^z \approx 0.000$, where a^z denotes the area power of the characteristics model, which means that about 100.0% of feasible outcomes are ruled out by the corresponding revealed preference restrictions. It is easy to see why this happens. An implication of the characteristics model is that the number of goods consumed cannot exceed the number of characteristics, unless characteristics budget sets are uniquely aligned. While the sets of feasible goods bundles in periods 1 and 2 include all points on the corresponding budget planes implied by observed prices and quantities, the set

of z-rationalizable goods bundles is restricted to points on the edges of these planes.[6] In general, in a model with fewer characteristics than goods, dimension reduction has the effect of increasing the area power by an infinite amount,[7] given a feasible outcome space corresponding to (P1).

In order to avoid bestowing any *a priori* advantage on the characteristics model, we next consider

$$\text{(P2)}\ \ \mathcal{F} = \times_{t=1}^{T}\mathcal{F}_t = \times_{t=1}^{T}\{q : p_t'q = p_t'q_t,\ p_t^0 = \infty,\ q \in \mathcal{R}_+^K\},$$

where p_t^0 is a sub-vector of p_t for which the corresponding $q_{kt} = 0$. This corresponds to a natural set of feasible outcomes when prices are missing. When unobserved prices are set arbitrarily high, we essentially are able to control for dimension reduction. Here we avoid assigning any weight to the increase in the area power of the characteristics model achieved purely by reducing the number of dimensions. Under (P2), $a^q = a^z = 1.000$ in the preceding example, which means that about 0.0% of feasible outcomes are ruled out by the revealed preference conditions for both the goods and characteristics models with missing prices. This happens because the feasible outcome space has been reduced to a single dimension, namely observed prices and quantities of goods purchased. In general, missing prices give rise to decreases in area power.[8]

For the set of feasible outcomes corresponding to the characteristics model, we also consider $\mathcal{F}_t = \{q : z_t R^D z \text{ and not } z_t P^D z\}$. Given a set of observations (p_t, q_t), $t = 1, \ldots, T$, and a technology A, z_t is directly revealed preferred to z, written $z_t R^D z$, if $p_t'q_t \geq p_t'q$ and there exist $\rho_t \in [1, \infty)$ and $\sigma_t \in \mathcal{R}^J$ such that $\rho_t p_t \geq A\sigma_t$, binding for any $q_{kt} > 0$. The transitive closure z_t revealed preferred to z, written $z_t R z$, as well as z_t strictly directly revealed preferred to z, written $z_t P^D z$, follow as expected. However, under these priors the set of feasible goods bundles in a given period may not contain the observed goods bundle. This happens because the characteristics bundle corresponding to the observed goods bundle is either in the interior or outside of the characteristics budget set. Since these anomalies are largely driven by imputing

[6] In this example, consuming some of good 1 and some of good 3 and consuming all goods are ruled out in both periods.

[7] Increases in area power correspond to smaller values of a^i for model $i \in \{q, z\}$ and vice versa.

[8] For the characteristics model, rather than setting unobserved prices arbitrarily high, it is sometimes possible to construct virtual prices that just allow for observed behavior. This necessarily increases the area power of the characteristics model.

unobserved prices, we opt in favor of allowing for missing prices.

An alternative strategy altogether is to discretize the feasible outcome space. Under (P1), dimension reduction has the effect of increasing area power for the characteristics model by an infinite amount only because the feasible outcome space is continuous, which is of course unrealistic in empirical practice. A solution is to assume that the feasible outcome space is discrete, which better reflects the practical nature of the choice problem and allows us to calculate exact areas.[9] Consider

(P3) $\mathcal{F} = \times_{t=1}^{T} \mathcal{F}_t = \times_{t=1}^{T} \{q : p_t' q = p_t' q_t,\ q \in \mathcal{Z}_+^K\}$.[10]

Note that in the limit areas calculated under (P3) tend towards those estimated under (P1). Under (P3), in the previous example there are 28 feasible goods bundles in each period if we restrict consumption to whole units, which gives 784 feasible outcomes, of which 619 are q-rationalizable, such that $a^q \approx 0.790$. Of these 784 feasible outcomes, only 106 are z-rationalizable, such that $a^z \approx 0.135$. If we further restrict consumption to half units in the preceding example, $a^q \approx 0.787$ and $a^z \approx 0.042$, tending towards the (P1) results. As we have shown in this section, the area power of a model heavily depends upon the feasible outcome space over which that model is defined. Once we have specified our feasibility priors, the remaining step in identifying the model which best fits the data is to make an axiomatic claim about predictive success.

2.5. *Predictive Success*

After defining area power as a set restriction, Beatty and Crawford (2009) make use of a theorem from Selten (1991), who has put forward a measure of predictive success for any theory that predicts a subset of all possible outcomes. This measure was developed initially for experimental game theory. For our purposes, the relative size of the predicted subset is a^i for model $i \in \{q, z\}$, and a set of observations is either an element of the predicted subset or it is not, which is captured by the pass/fail indicator $r^i \in \{0, 1\}$. Loosely speaking, a^i represents the size of the target, and r^i a

[9]The revealed preference conditions for the goods and characteristics models do not depend upon assuming a continuous outcome space. As long as discrete consumption is weakly separable from continuous consumption, the familiar results hold (see Polisson and Quah (2011) for a formal treatment of discreteness and revealed preference).

[10]Here we have restricted consumption to whole units, but discreteness can be defined more generally.

hit or a miss of the target. In order to assess the performance of an area theory, we require a predictive success measure $m(r^i, a^i) : \{0, 1\} \times [0, 1] \to \mathcal{R}$. That $m(\cdot)$ takes as its arguments (r^i, a^i) is relatively uncontroversial, but the functional form for $m(\cdot)$ is entirely debatable. Selten (1991) suggests several desirable properties.

DEFINITION 2.3 A predictive success measure $m(r^i, a^i)$ satisfies

1. *Monotonicity* if $m(1, 0) > m(0, 1)$,
2. *Equivalence* if $m(0, 0) = m(1, 1)$, and
3. *Aggregability* if $m(\lambda r_1^i + (1-\lambda) r_2^i, \lambda a_1^i + (1-\lambda) a_2^i) = \lambda m(r_1^i, a_1^i) + (1-\lambda) m(r_2^i, a_2^i)$ for any $\lambda \in [0, 1]$.

Monotonicity states that passing infinitely stringent restrictions is a greater success than failing infinitely lenient restrictions; equivalence states that failing infinitely stringent restrictions and passing infinitely lenient restrictions are equally informative; and aggregability states that a predictive success measure should allow for additive decomposition such that it can be applied easily to a heterogenous sample. Given these axioms, Selten (1991) derives the following result.

THEOREM 2.3 *The predictive success measure* $m(r^i, a^i) = r^i - a^i \in [-1, 1]$ *satisfies monotonicity, equivalence, and aggregability. If another predictive success measure* $\tilde{m}(r^i, a^i)$ *also satisfies these axioms, then there exist* $\alpha^i \in \mathcal{R}$ *and* $\beta^i \in \mathcal{R}_{++}$ *such that* $\tilde{m}(r^i, a^i) = \alpha^i + \beta^i m(r^i, a^i)$.

PROOF: See Selten (1991) and Beatty and Crawford (2009). *Q.E.D.*

The main implication of the theorem is not only that $m(r^i, a^i) = r^i - a^i$ satisfies three desirable axioms, but also that any other measure satisfying these axioms is an affine transformation of $r^i - a^i$. This result suggests that $m(r^i, a^i) = r^i - a^i$ is a suitable measure for establishing predictive success. The higher this measure, the more successful the model, and vice versa. When $r^i - a^i$ is close to zero, the set of observations is either passing lenient restrictions of failing stringent restrictions.

From Theorem 2.3, we know that if $m(r^z, a^z) > m(r^q, a^q)$, then the characteristics model outperforms the goods model, and vice versa. If $m(r^z, a^z) = m(r^q, a^q)$, then the two models are equally predictively successful. This suggests that differencing the two measures is informative about *relative* predictive success. Selten (1991) derives

an ordinal counterpart to the cardinal result in Theorem 2.3, which depends upon an axiom stating that any two theories predicting subsets of all possible outcomes should be compared on the basis of changes in accuracy and precision. Within our framework, $\Delta r = r^z - r^q$ denotes the change in accuracy and $\Delta a = a^z - a^q$ the change in precision. It can be shown that $\delta(\Delta r, \Delta a) = \Delta r - \Delta a = m(r^z, a^z) - m(r^q, a^q)$ is a suitable measure for establishing relative predictive success (see Appendix A.4 for an axiomatization of relative predictive success).[11] The higher this measure, the more relatively successful the characteristics model, and vice versa. When $\Delta r - \Delta a$ is close to zero, the goods and characteristics models perform equally well. Note that this result applies to area theories more generally.

2.6. *Extrapolation*

The revealed preference procedures outlined in previous sections are easily applied to observed data, with an aim to relating consistency and power to establish fit. In order to explore comparative statics, we introduce hypothetical variation and then test for further consistency. As Varian (1982) shows for the goods model, this amounts to introducing new price vectors and testing for q-rationalizability, which then establishes nonparametric revealed preference bounds on demand responses, or set-identification. A series of papers by Blundell, Browning, and Crawford (2003, 2007, 2008) introduces methods to tighten nonparametric revealed preference bounds, either by exploiting income expansion paths or separability in preferences. We follow in the tradition of the latter innovation. For each set of observations, we extrapolate according to the model which best fits the data. This allows us to tighten bounds on demand responses when the set of observations is consistent with the characteristics model.

We illustrate with a simple simulation, in which we assume Cobb-Douglas preferences over characteristics. As shown in Figure 4, the characteristics bounds on own-price demand responses are weakly tighter than the goods bounds, which follows since z-rationalization implies q-rationalization (see Appendix A.5 for a formal

[11] This is a subtle point. Relative predictive success might well be captured by some measure like an elasticity, but we find that a simple differencing of predictive success measures has the requisite properties.

treatment of nested bounds).

[Figure 4 about here]

The true demand function for a particular good is represented by the solid line and the respective bounds by the dotted lines in Figures 4a and 4b. Given this particular underlying characteristics structure, the own-price demand response is discontinuous, i.e., the individual exhausts her income on the good up to a threshold price, after which she switches to another good or goods. This knife-edge feature is common to characteristics models. As shown in Figure 4a, the goods bounds are in general very weak. The upper bound traces exhaustion, and the lower bound is zero. However, the demand function is point-identified at the observed prices. As shown in Figure 4b, the characteristics bounds are better than the goods bounds. We obtain the same point identification as in the goods case, along with improvements to both the upper and lower bounds. Furthermore, as shown in Figure 5, the characteristics bounds on cross-price demand responses are weakly tighter than the goods bounds as well.

[Figure 5 about here]

These simulations are extremely simple, with minimal goods, characteristics, and time periods. Since the power of the revealed preference conditions is improved by adding data, we could always hope to improve these bounds by observing consumption over more periods. Nonetheless, we highlight that extrapolation allows us to draw inference about theoretically consistent substitution patterns, and that this inference is made stronger by assuming additional structure in the data-generation process.

3. IMPLEMENTATION

3.1. *Application*

We implement the procedures outlined in previous sections on household panel data from the UK milk market.[12] We choose milk to establish continuity with an existing literature (see, for example, Blow, Browning, and Crawford (2008) and Griffith and Nesheim (2010)) and to highlight a relevant nutritional application. The procedures developed in this paper allow us to identify the relative importance of nutrients

[12]Programs are available as part of a supplementary web appendix.

(characteristics) in explaining food (goods) consumption.[13] We envisage a model in which individuals trade off between taste and health, which is of particular interest for goods that we may perceive to ultimately become "bads". The relevant policy interest may be whether an effective tax on fat, for example, manifested as a tax on particular foods, causes people to eat more healthily.

3.2. *Data*

Our data are from the Kantar UK Worldpanel, which is a representative rolling panel of more than 15,000 households since 2001 and more than 25,000 households since 2006. It contains standard demographic information for each household and its constituent members. It also contains detailed expenditure and product information for fast-moving consumer goods since 2001, including nutritional information since 2006. Each household is equipped with a barcode scanner, which records the precise details associated with each purchase. These data are are then linked to manufacturer product databases. The relative strengths of the Kantar data (i.e., in comparison with traditional sources like the Expenditure and Food Survey (EFS), formerly the Family Expenditure Survey (FES)) are that they follow households over time and that they provide a high level of detail. The relative weaknesses come primarily as a consequence of respondent fatigue. See Griffith and O'Connell (2009) and Leicester and Oldfield (2009) for further descriptions of the Kantar data.

We follow a sample of UK households for a full year from November 1, 2006 to October 31, 2007. We focus on the three main types of conventional non-organic milk: whole, semi-skimmed, and skimmed.[14] Following Blow, Browning, and Crawford (2008), we aggregate household milk purchases to a monthly level in order to reduce the computational burden and to respect the intertemporal separability implicit in the models we are testing.[15] For each household and month, we observe quantities and

[13]Since food contains some nutrients that are healthy in small amounts and unhealthy in large amounts, the characteristics model is a very plausible structure for food and nutrients.

[14]We exclude all organic, soya, UHT, speciality milks (e.g., buttermilk, chocolate milk, goats' milk, etc.), and super fat milks. We also exclude promotional purchases, as well as cases with pack size discrepancies and expenditure anomalies. We then focus on the major brands, each of which has at least a 1% market share by volume in our sample.

[15]Since milk must be consumed quickly, aggregating by month effectively treats milk as a non-durable, non-storable good. Note, as in Blow, Browning, and Crawford (2008), that aggregation is

expenditures for each type of milk that is purchased, from which we construct unit prices. For types that are not purchased, we impute missing unit prices by taking the medians of observed unit prices by region and month.[16] We round monthly quantities to the nearest pint and unit prices to the nearest pence. Since most milk is sold in units of pints in England, 91% of the market share by volume in our sample, rounding has only a negligible effect on monthly expenditures. Finally, we extract a relatively homogenous subsample of couples with children, which gives us a working sample of 5,648 households.

Table I shows market shares by value and volume, and Table II displays unit prices (£/pint) for each type of milk.[17]

[Tables 1 and 2 about here]

Semi-skimmed milk has more than 55% of the market share by both value and volume, followed by whole milk at about 35% and skimmed milk at about 8%. In comparison to a Danish sample (see Blow, Browning, and Crawford (2008)), this sample of Britons drink significantly more whole milk, slightly less semi-skimmed milk, and significantly less skimmed milk on average. Furthermore, we do not observe a similar price gradient with respect to fat content as in Denmark. Whole and semi-skimmed milk are about £0.31/pint, and skimmed milk is slightly more expensive at about £0.32/pint.

Note that implicit in our implementation is the assumption that milk characteristics are weakly separable from all other characteristics. This is perhaps contentious, but it is also necessary in order to make the implementation tractable. However, we note that the procedures developed in this paper can be extended to more complicated systems, which is of particular value in the nutritional context mentioned previously, where the effects of hypothetical own- and cross-price variation and any corresponding substitution patterns are of significant academic and policy interest. We leave this as a subject of future research.

Our application has $K = 3$ goods and $J = 2$ characteristics, and the technology A

a simplification that potentially obscures container size and other relevant priced characteristics.

[16]Kantar partitions the UK into 10 geographical regions: London, Midlands, North East, Yorkshire, Lancashire, South, Scotland, Anglia, Wales and West, and South West. Imputing missing unit prices at a much finer regional level, i.e., postcode, is not always possible in the data.

[17]Note that these figures are based on observed data only.

is given by

$$A = \begin{pmatrix} 1 & 20.869 \\ 1 & 9.429 \\ 1 & 0.568 \end{pmatrix}. \tag{3.1}$$

The three goods are whole, semi-skimmed, and skimmed milk. Following Blow, Browning, and Crawford (2008), the first characteristic is "milkiness", which is the combination of water, calcium, vitamins, etc., that can be found in any milk. A pint of milk contains a pint of milkiness. The second characteristic is fat (grams/pint), which of course varies across types of milk. A pint of whole, semi-skimmed, and skimmed milk contains 20.869, 9.429, and 0.568 grams of fat on average by volume, respectively. Therefore, consuming a pint of each type of milk provides 30.866 grams of fat altogether. Note that since preferences over milkiness and fat are non-satiated, the structure allows fat to be a "bad" characteristic.

4. RESULTS

The consistency results are shown in Table III.

[Table 3 about here]

Here we find that of 5,648 households, about 90.0% are q-rationalizable and 52.0% are z-rationalizable with imputed prices. As expected, the pass rate for the goods model is higher than the pass rate for the characteristics model since z-rationalization implies q-rationalization. Similarly, about 97.7% of households are q-rationalizable and 88.3% are z-rationalizable with missing prices. Recall that the revealed preference conditions with missing prices are more permissive than with imputed prices. Our results are in line with the consistency results in Blow, Browning, and Crawford (2008).

The major contribution of this paper lies in identifying the model which best fits the data in light of these pass rates, and in order to do so, we consider the area power of the goods and characteristics models. As shown in Table IV, averaged across all households, $a^q \approx 0.857$ and $a^z \approx 0.000$ under (P1).

[Table 4 about here]

Recall that dimension reduction often has the effect of reducing characteristics areas to zero under (P1). Consequently, averaged across all households, $m(r^q, a^q) \approx 0.043$

and $m(r^z, a^z) \approx 0.520$ under (P1). For the goods model, this figure masks a great deal of heterogeneity. Among 5,648 households, 564 (10.0%) have negative predictive success, 2,750 (48.7%) have zero predictive success, and 2,334 (41.3%) have positive predictive success. The standard deviation of predictive success is 0.291. For the characteristics model, predictive success is completely determined by the pass rate. Lastly, averaged across all households, $\delta(\Delta r, \Delta a) \approx 0.477$ under (P1). Once again, this figure obscures some important heterogeneity. Among 5,648 households, 1,203 (21.3%) have negative relative predictive success, 952 (16.9%) have zero relative predictive success, and 3,493 (61.8%) have positive relative predictive success. The standard deviation of relative predictive success is 0.538. While the sample, on average, favors the characteristics structure, a sizable 38.2% of households favor the goods structure or indifference. Since these procedures are maximally heterogeneous, we later some explore logical associations with observable household attributes that may be useful in explaining some of this variation.

As previously noted, the tremendous area power advantage of the characteristics model relative to the goods model under (P1) is largely attributable to the dimension-reducing properties of the former structure. Under (P2), unobserved prices are set arbitrarily high, which restricts the number of feasible dimensions at the outset and therefore effectively controls for dimension reduction. As shown in Table IV, averaged across all households, $a^q \approx 0.948$ and $a^z \approx 0.861$ under (P2), which implies that $a^z - a^q \approx -0.087$ is the average change in precision attributable to the characteristics structure alone. However, the average change in accuracy $r^z - r^q \approx -0.094$ under missing prices is slightly greater than the average change in precision, which means that averaged across all households, $\delta(\Delta r, \Delta a) \approx -0.007$ under (P2). After controlling for dimension reduction, there is little between the two models, and even a slightly greater support for the goods structure on average.[18] This result suggests that in this instance much of the increase in area power associated with the charac-

[18]However, recall that for the characteristics model, it is sometimes possible to construct virtual prices that just allow for observed behavior. This procedure also controls for dimension reduction, but just, and therefore necessarily increases the area power of the characteristics model. As a result, 0.861 is an upper bound on a^z averaged across households, 0.022 is a lower bound on $m(r^z, a^z)$ averaged across households, and -0.007 is a lower bound on $\delta(\Delta r, \Delta a)$ averaged across households under (P2).

teristics structure is attributable to its dimension-reducing capabilities rather than the technological mapping.

As before, the (P2) results in Table IV mask some important heterogeneity. For the goods/characteristics models, among 5,648 households, 131/104 (2.3%/1.8%) have negative predictive success, 4,764/5,015 (84.3%/88.8%) have zero predictive success, and 753/529 (13.3%/9.4%) have positive predictive success, respectively. The standard deviation of predictive success is 0.175/0.153 in the goods/characteristics models, respectively. Furthermore, averaged across all households, $\delta(\Delta r, \Delta a) \approx -0.007$ also obscures some heterogeneity. Among 5,648 households, 528 (9.3%) have negative relative predictive success, 4,799 (85.0%) have zero relative predictive success, and 321 (5.7%) have positive relative predictive success. The standard deviation of relative predictive success is 0.133. While the sample, on average, favors the goods structure very slightly, a sizable 90.7% of households favor the characteristics structure or indifference.

Dimension reduction has a pronounced effect on area power in part because the outcome space is continuous. Under (P3), a discrete consumption space moderates these effects. As shown in Table IV, averaged across all households, $a^q \approx 0.743$ and $a^z \approx 0.134$ under (P3). Consequently, averaged across all households, $m(r^q, a^q) \approx 0.157$ and $m(r^z, a^z) \approx 0.386$ under (P3). For the goods/characteristics models, among 5,648 households, 440/181 (7.8%/3.2%) have negative predictive success, 2,504/3,029 (44.3%/53.6%) have zero predictive success, and 2,704/2,438 (47.9%/43.2%) have positive predictive success. The standard deviation of predictive success is 0.322/0.476 in the goods/characteristics models, respectively. Lastly, averaged across all households, $\delta(\Delta r, \Delta a) \approx 0.229$ under (P3). Among 5,648 households, 1,463 (25.9%) have negative relative predictive success, 1,332 (23.6%) have zero relative predictive success, and 2,853 (50.5%) have positive relative predictive success. The standard deviation of relative predictive success is 0.529. While the sample, on average, favors the characteristics structure, a sizable 49.5% of households favor the goods structure or indifference.

Altogether, the results suggest that the characteristics model fits the data better than does the goods model, but that dimension reduction rather than the characteristics structure is the determining factor. The relative predictive success measure

shows that the characteristics model performs relatively best under (P1), followed by (P3) and then (P2). The (P1) results very strongly support the characteristics structure, which is largely due to dimension reduction over a continuous outcome space. In order to mitigate against this *a priori* favoritism, we control for dimension reduction under (P2) and find there to be very little between the two models. Lastly, we moderate the effects of dimension reduction by discretizing the outcome space under (P3). One of our most important messages is that any claims about power and fit depend heavily upon feasibility priors, and also that it is important to isolate the changes in power and fit associated with dimension reduction from those associated with the particular structure of the model. In our example, had we observed positive relative predictive success under (P2), we might conclude that both dimension reduction *and* the characteristics structure are important in explaining some of the observed variations in the data. However, in this instance, alternatives such as weakly separable or other characteristic structures may have performed better.

Here we briefly investigate via a descriptive regression analysis whether consistency, power, and fit are associated with observable household attributes. The controls include dummy variables corresponding to income bands and the main shopper being female, as well as continuous variables corresponding to the age of the main shopper in the household and the size of the household. Overall the statistical explanatory power of these control variables is incredibly weak, as suggested by low R^2 and pseudo-R^2 values, which leads us to conclude that consistency, power, and fit are largely unexplained by observable household attributes. However, we observe a slight income gradient, whereby wealthier households are less likely to be consistent with either model, which disappears in the goods case but not the characteristics case after accounting for area power. Furthermore, households with older main shoppers are more likely to be consistent with either model, but these effects are minimized after adjusting for area power. The female main shopper dummy variable fails to give a significant result across many specifications. Lastly, the effect of household size has a positive and statistically significant effect on consistency and predictive success for the characteristics but not the goods model, and consequently, on relative predictive success for the characteristics model.

Finally, we demonstrate how to explore comparative statics via set-identification.

To do this, we randomly select a household that is consistent with both the goods and characteristics models, but which favors the characteristics model, i.e., which exhibits positive relative predictive success. We observe

$$q = \begin{pmatrix} 0 & 0 & 0 \\ 4 & 4 & 4 \\ 0 & 0 & 0 \end{pmatrix} \text{ and } p = \begin{pmatrix} 0.28 & 0.28 & 0.29 \\ 0.28 & 0.27 & 0.30 \\ 0.32 & 0.32 & 0.33 \end{pmatrix}, \tag{4.1}$$

where the rows of quantities q and prices p correspond to goods and the columns to periods, and where the prices of goods not purchased are imputed. Here we find that $a^q \approx 0.602$ and $a^z \approx 0.102$ under (P3), and since the set of observations is consistent with both models, this gives $m(r^q, a^q) \approx 0.398$, $m(r^z, a^z) \approx 0.898$, and $\delta(\Delta r, \Delta a) \approx 0.500$. The characteristics model outperforms the goods models for this particular household by our criteria, which allows for a tightening of the bounds on own- and cross-price demand responses.

As illustrated in Figure 6, the characteristics bounds on own-price demand responses are better than the goods bounds, both evaluated at median prices and expenditures.

[Figure 6 about here]

The improvement comes primarily from a tightening of the upper bound. Similarly, as shown in Figure 7, the characteristics bounds on cross-price demand responses are better than the goods bounds, again evaluated at median prices and expenditures, and these improvements are more pronounced.

[Figure 7 about here]

Altogether these results suggest that imposing a characteristics structure, when appropriate, may be useful in identifying substitution patterns. However, a limitation here is that in order to make use of this information, we might want to aggregate across households. We leave this as a subject of future research.

5. CONCLUSIONS

In this paper, we have compared the goods and characteristics models of the consumer within a traditional demand framework. We have examined the nonparametric

revealed preference conditions for the goods and characteristics models, and we have developed a methodology for testing nested models of this class using nonparametric revealed preference techniques. Of primary interest has been to make a comparison on the basis of predictive success, which has required that we develop a method to relate set predictions across models. This has allowed us to nonparametrically identify the model which best fits the data, and in doing so, to identify the value added by the characteristics structure in explaining consumer behavior. We have then explored the effects of hypothetical price variation as implied by our findings in order to nonparametrically bound comparative statics of interest.

In the implementation on household panel data from the UK milk market, we have shown that the characteristics model fits the data better than does the goods model, and in such cases, we can improve bounds on own- and cross-price demand responses. Most importantly, we have shown that the better fit of the characteristics model is entirely attributable to dimension reduction. As a caution, we note that since the restrictive power of a model depends upon the feasible outcome space over which it is defined, any results must be interpreted in light of their corresponding feasibility priors. An analogy in statistical testing is that the result of the test depends upon the null hypothesis to be rejected. This should come as no surprise to econometricians, who are in the habit of constructing hypotheses in light of the possibility of Type I and Type II errors.

The methods developed in this paper can, in principle, be applied to other models of a similar class, where we can exploit structure/power differentials to compare models on the basis of predictive success. Since the dimension-reducing property of the characteristics structure is the driving force behind its improved fit, there may be another dimension-reducing structure that provides an even better fit. Applied microeconometricians have long recognized the value of dimension reduction in empirical work. It simplifies the problem, both for the agent and the observer, and at the very least is implicitly invoked in nearly all applied empirical work. Typically, separability restrictions or characteristics structures are introduced to reduce the number of dimensions, but it is *a priori* unclear how or where these restrictions should be imposed. Future research on dimension reduction shall explore where it pays dividends in terms of fit to be strict versus lenient with structural assumptions.

In the implementation, we have highlighted a relevant nutritional application. The procedures developed in this paper have allowed us to nonparametrically identify the relative importance of nutrients in explaining food consumption. Economists are often interested in estimating theoretically consistent demand systems for food. Since food contains some nutrients that are healthy in small amounts and unhealthy in large amounts, the characteristics model is a very plausible structure for food and nutrients. Individuals may trade off between taste and health, which is of particular interest for goods that we may perceive to ultimately become "bads". The procedures developed in this paper can be extended to more complicated demand systems, which is of particular value in this nutritional context, where the effects of hypothetical own- and cross-price variation and corresponding substitution patterns are of significant academic and policy interest. If consumers have preferences over nutrients rather than food, then the implications of price variation change dramatically. We leave this as a subject of future research. There are of course other markets to explore, e.g., housing, automobile, education, and labor, which may be of great interest to applied researchers and policymakers, which again we leave for future work.

This paper has neglected any considerations on the supply side of the economy. However, applied microeconometricians working on problems of industrial organization are very often interested in both sides of the market. In the context of the characteristics framework, prices are endogenous, and hence there is an equilibrium model underlying hedonic prices. The basic equilibrium theory posits that individual consumers and differentiated producers make choices in characteristics space, the results of which are market-clearing equilibrium prices that reflect implicit intrinsic mappings from the prices of goods to the characteristics they embody. In order to apply the results of this paper to applied microeconomic work in these areas, we must further explore the driving forces behind identification in equilibrium hedonic models under minimal assumptions.

APPENDIX A: THEORY APPENDIX

A.1. *Empirical Implications of the Characteristics Model*

The primal problem in the linear characteristics model of Section 2.1 is to

$$\max_q v(A'q) \text{ subject to } p'q \leq x. \tag{A.1}$$

Here we note that where $A \in \mathcal{R}_+^{K \times J}$, $J < K$ and A has full rank by construction. Lancaster (1966) discuses three alternative technological structures: (1) when $J = K$, the characteristics model collapses to a goods model with transformed prices since $q = (A')^{-1}z$ as long as A has full rank; (2) when $J > K$, $z = A'q$ contains more equations than unknowns, such that a q may not exist for every z; and (3) when $J < K$, $z = A'q$ contains fewer equations than unknowns, such that an infinite number of q may exist for any z. This third non-trivial scenario is most economically interesting as it reduces the dimensions of the choice problem for the consumer. The problem in (A.1) gives rise to the first-order conditions $p \geq A\pi$, binding for any $q_k > 0$, where $\pi = Dv(z)/\lambda$ are the shadow prices of characteristics, which are linearly combined to form the shadow prices of goods. The result is the demand system $q = \tilde{q}(p, x, A)$. The dual problem is to

$$\min_q p'q \text{ subject to } v(A'q) \geq v, \tag{A.2}$$

which has similar first-order conditions to those in the primal problem and gives the Hicksian demand system $q = \tilde{h}(p, v, A)$. Using $\tilde{q}(\cdot)$ and $\tilde{h}(\cdot)$, we construct the indirect utility and expenditure functions, respectively, and we exploit the usual duality results. It can be shown that $\tilde{q}(\cdot)$ is differentiable and zero homogenous in (p, x), satisfies adding-up, and gives rise to a Slutsky matrix that is symmetric and negative semi-definite. However, there is a further restriction that is best illustrated diagrammatically with three goods and two characteristics. In the example in Figure 8a, the characteristics budget set restricts the individual to consuming only good 1, only good 2, only good 3, some of good 1 and some of good 2, or some of good 2 and some of good 3.

[Figure 8 about here]

Consuming some of good 1 and some of good 3 and consuming all goods are both ruled out in this particular scenario. One can imagine other characteristics budget sets that give rise to similar dimension-reducing restrictions. In the example in Figure 8b, the individual is free to consume in any combination. These restrictions are neatly encapsulated by the rank condition in Section 2.1, which states that the number of goods consumed cannot exceed the number of characteristics, unless characteristics budget sets are uniquely aligned as in Figure 8b.

A.2. *Empirical Implications of Nested Models*

While we focus on the characteristics model in this paper, nested restrictions enable us to identify the relative importance of additional structure more generally. For example, without assuming that preferences take a particular functional form, separability is a structural imposition that reflects natural groupings and is often introduced to simplify the consumer's problem. The very general non-separable goods model nests any additional separable structure that we may wish to impose, and again as a consequence, any further empirical restrictions are also nested. The spectrum typically runs from weak separability, which partitions goods into mutually exclusive groupings, to latent separability (see Blundell and Robin (2000)), which allows for overlapping groupings, to characteristics, which maps from goods to the characteristics they embody. The implications of weak separability

have of course produced a long and rich literature. Early work by Leontief (1947a,b) can be used to show that weak separability is a necessary and sufficient condition for within-group demands to depend only upon within-group prices and group expenditures, a convenient result that is at the very least invoked implicitly in nearly all empirical work. Weak separability makes possible the two-stage budgeting procedures of Strotz (1957) and Gorman (1959), where within-group price aggregation further requires that group sub-utility functions are homothetic. As a final example, the collective model of the household (see, for example, Chiappori (1988, 1992) and Browning and Chiappori (1998)) nests the unitary model.

A.3. *Revealed Preference Definitions, Axioms, and Theorems*

Let $q_t \in \mathcal{R}^K_+$ be observed goods bought at prices $p_t \in \mathcal{R}^K_{++}$ in period $t \in \{1, \ldots, T\}$.

DEFINITION A.1 Given a set of observations (p_t, q_t), $t = 1, \ldots, T$,

1. q_t is *directly revealed preferred* to q_s, written $q_t R^D q_s$, if $p_t' q_t \geq p_t' q_s$,
2. q_t is *revealed preferred* to q_s, written $q_t R q_s$, if $p_t' q_t \geq p_t' q_r$, $p_r' q_r \geq p_r' q_n$, $\ldots$, $p_m' q_m \geq p_m q_s$, and
3. q_t is *strictly directly revealed preferred* to q_s, written $q_t P^D q_s$, if $p_t' q_t > p_t' q_s$.

DEFINITION A.2 A set of observations (p_t, q_t), $t = 1, \ldots, T$, satisfies

1. the *Weak Axiom of Revealed Preference (WARP)* if $q_t R^D q_s$ implies not $q_s R^D q_t$ $\forall$ $s, t = 1, \ldots, T$,
2. the *Strong Axiom of Revealed Preference (SARP)* if $q_t R q_s$ and $q_t \neq q_s$ together imply not $q_s R q_t$ $\forall$ $s, t = 1, \ldots, T$, and
3. the *Generalized Axiom of Revealed Preference (GARP)* if $q_t R q_s$ implies not $q_s P^D q_t$ $\forall$ $s, t = 1, \ldots, T$.

Note that, unlike WARP, SARP and GARP both utilize the transitive closure, and also that GARP allows for demand correspondences while SARP only allows for single-valued demand functions.

If the set of observations (p_t, q_t), $t = 1, \ldots, T$, is consistent with the goods model of the consumer, it can be rationalized as follows:

DEFINITION A.3 A utility function $u(\cdot) : \mathcal{R}^K_+ \to \mathcal{R}$ q-rationalizes the set of observations (p_t, q_t), $t = 1, \ldots, T$, if $u(q_t) \geq u(q)$ for all q such that $p_t' q_t \geq p_t' q$.

The necessary and sufficient conditions for q-rationalization and GARP are related by the full version of Afriat's Theorem (see Afriat (1967)) as follows:

THEOREM A.1 *The following statements are equivalent:*

1. *There exists a utility function* $u(\cdot) : \mathcal{R}^K_+ \to \mathcal{R}$ *that is non-satiated which q-rationalizes the set of observations* (p_t, q_t), $t = 1, \ldots, T$.

2. *The set of observations* (p_t, q_t), $t = 1, \ldots, T$, *satisfies GARP.*

3. *There exist* $U_t \in \mathcal{R}$ *and* $\lambda_t \in \mathcal{R}_{++}$, $t = 1, \ldots, T$, *such that*

$$U_s \leq U_t + \lambda_t p_t'(q_s - q_t) \ \forall \ s, t = 1, \ldots, T.$$

4. *There exists a utility function* $u(\cdot) : \mathcal{R}^K_+ \to \mathcal{R}$ *that is continuous, increasing, and concave which q-rationalizes the set of observations* (p_t, q_t), $t = 1, \ldots, T$.

PROOF: See Afriat (1967), Varian (1982, 1983), and Fostel, Scarf, and Todd (2004). *Q.E.D.*

The main implication of the theorem is that if there exists a solution to the system of T^2 linear inequalities and $2T$ unknowns in statement (3), then the set of observations is q-rationalizable, or consistent with the goods model of the consumer. While this is computationally viable for small dimensional problems, it becomes increasingly difficult for large dimensional problems. An alternative strategy is to test for violations of GARP, which is computationally very rapid. Further note that the implication of statements (1) and (4) is that non-satiation versus continuity, monotonicity, and concavity are indistinguishable within a finite set of observations.

Let $z_t \in \mathcal{R}^J_+$ be observed characteristics bought at prices $p_t \in \mathcal{R}^K_{++}$ in period $t \in \{1, \ldots, T\}$, and let $A \in \mathcal{R}^{K \times J}_+$ be a linear technology mapping from q_t to z_t according to $z_t = A'q_t \ \forall \ t = 1, \ldots, T$. If the set of observations (p_t, q_t), $t = 1, \ldots, T$, and the technology A are consistent with the linear characteristics model of the consumer, they can be rationalized as follows:

DEFINITION A.4 A utility function $v(\cdot) : \mathcal{R}^J_+ \to \mathcal{R}$ z-rationalizes the set of observations (p_t, q_t), $t = 1, \ldots, T$, for the technology A if $v(z_t) = v(A'q_t) \geq v(A'q) = v(z)$ for all q such that $p_t'q_t \geq p_t'q$.

The necessary and sufficient conditions for z-rationalization and GARP are related in full by Blow, Browning, and Crawford (2008) as follows:

THEOREM A.2 *The following statements are equivalent:*

1. *There exists a utility function* $v(\cdot) : \mathcal{R}^J_+ \to \mathcal{R}$ *that is non-satiated which z-rationalizes the set of observations* (p_t, q_t), $t = 1, \ldots, T$, *for the technology A.*

2. *The set of observations* $(\pi_t, A'q_t)$, $t = 1, \ldots, T$, *satisfies GARP for the technology A and some* $\pi_t \in \mathcal{R}^J$, $t = 1, \ldots, T$, *such that*

$$p_t \geq A\pi_t \ \forall \ t = 1, \ldots, T,$$

binding for any $q_{kt} > 0$.

3. *There exist* $V_t \in \mathcal{R}$, $\lambda_t \in \mathcal{R}_{++}$, *and* $\pi_t \in \mathcal{R}^J$, $t = 1, \ldots, T$, *such that*

$$V_s \leq V_t + \lambda_t \pi_t'(A'q_s - A'q_t) \text{ and}$$
$$p_t \geq A\pi_t \ \forall \ s, t = 1, \ldots, T,$$

the latter binding for any $q_{kt} > 0$.

4. *There exist* $U_t \in \mathcal{R}$, $\rho_t \in [1, \infty)$, *and* $\sigma_t \in \mathcal{R}^J$, $t = 1, \ldots, T$, *such that*

$$U_s \leq U_t + \sigma_t'(A'q_s - A'q_t),$$
$$\rho_t p_t \geq A\sigma_t \ \forall \ s, t = 1, \ldots, T,$$

the latter binding for any $q_{kt} > 0$.

5. *There exists a utility function* $v(\cdot) : \mathcal{R}^J_+ \to \mathcal{R}$ *that is continuous, non-satiated, and concave which* z*-rationalizes the set of observations* (p_t, q_t), $t = 1, \ldots, T$, *for the technology* A.

PROOF: See Blow, Browning, and Crawford (2008). *Q.E.D.*

The main implication of the theorem is that if there exists a solution to the $T(K + T)$ linear (in)equalities and $T(J+2)$ unknowns in statement (4), then the set of observations is z-rationalizable, or consistent with the characteristics model of the consumer. Note that statement (4) is a derived counterpart to statement (3), introduced to facilitate testing by eliminating nonlinearities in the constraints. Further note that the GARP condition in statement (2) is not easily testable since the shadow prices of characteristics are unobserved. Lastly, note as before that the implication of statements (1) and (5) is that non-satiation versus continuity, non-satiation, and concavity are indistinguishable within a finite set of observations.

In order to reduce the computational burden of a large dimensional problem, it is sometimes useful to apply the necessary rank condition for z-rationalization, which states that

$$\text{(A.3)} \qquad rank(A_t^+ \sim p_t^+) = rank(A_t^+).$$

Note that if the rank condition is violated in a single period, then the set of observations is not z-rationalizable. In contrast, we require at least two periods to establish q-rationalizability.

A.4. *Axiomatization of Relative Predictive Success*

In order to assess the relative performance of a model, we require a relative predictive success measure $\delta(\Delta r, \Delta a) : \{-1, 0, 1\} \times [-1, 1] \to \mathcal{R}$.[19] That $\delta(\cdot)$ takes as its arguments $(\Delta r, \Delta a)$ and not (r^1, r^2, a^1, a^2) is not uncontroversial. Following Selten (1991), we require that comparisons of relative performance are location-independent, and we suggest the following desirable properties.

DEFINITION A.5 A relative predictive success measure $\delta(\Delta r, \Delta a)$ satisfies

[19] When comparing theories 1 and 2, $\Delta r = r^2 - r^1 \in \{-1, 0, 1\}$ and $\Delta a = a^2 - a^1 \in [-1, 1]$.

1. *Ordinality* if $\delta(1,-1) > \delta(0,-1) = \delta(1,0) > \delta(-1,-1) = \delta(0,0) = \delta(1,1) > \delta(-1,0) = \delta(0,1) > \delta(-1,1)$, and
2. *Aggregability* if $\delta(\lambda\Delta r_1+(1-\lambda)\Delta r_2, \lambda\Delta a_1+(1-\lambda)\Delta a_2) = \lambda\delta(\Delta r_1,\Delta a_1)+(1-\lambda)\delta(\Delta r_2,\Delta a_2)$ for any $\lambda \in [0,1]$.

Monotonicity/equivalence establishes ordinal positions for different combinations of changes in accuracy and precision, and aggregability again allows for additive decomposition. Given these axioms, we derive the following result.

THEOREM A.3 *The relative predictive success measure* $\delta(\Delta r,\Delta a) = \Delta r - \Delta a \in [-2,2]$ *satisfies ordinality and aggregability. If another relative predictive success measure* $\tilde{\delta}(\Delta r,\Delta a)$ *also satisfies these axioms, then there exist* $\alpha \in \mathcal{R}$ *and* $\beta \in \mathcal{R}_{++}$ *such that* $\tilde{\delta}(\Delta r,\Delta a) = \alpha + \beta\delta(\Delta r,\Delta a)$.

Once again, the main implication of the theorem is not only that the relative predictive success measure $\delta(\Delta r,\Delta a) = \Delta r - \Delta a$ satisfies desirable axioms, but also that any other measure satisfying these axioms is an affine transformation of $\Delta r - \Delta a$. This result suggests that $\delta(\Delta r,\Delta a) = \Delta r - \Delta a$ is a suitable measure for establishing relative predictive success.

PROOF: This proof follows Selten (1991).[20] By aggregability,

$$\tilde{\delta}\left(-\frac{1}{2},-\frac{1}{2}\right) = \frac{1}{2}\tilde{\delta}(-1,-1) + \frac{1}{2}\tilde{\delta}(0,0), \tag{A.4}$$

and therefore by ordinality,

$$\tilde{\delta}\left(-\frac{1}{2},-\frac{1}{2}\right) = \tilde{\delta}(-1,-1) = \tilde{\delta}(0,0) = \tilde{\delta}(1,1). \tag{A.5}$$

Again by aggregability,

$$\tilde{\delta}\left(-\frac{1}{2},-\frac{1}{2}\right) = \frac{1}{2}\tilde{\delta}(0,-1) + \frac{1}{2}\tilde{\delta}(-1,0), \tag{A.6}$$

which together with (A.5) gives

$$\tilde{\delta}(-1,-1) = \tilde{\delta}(0,0) = \tilde{\delta}(1,1) = \frac{1}{2}\tilde{\delta}(0,-1) + \frac{1}{2}\tilde{\delta}(-1,0). \tag{A.7}$$

By similar arguments, it can be shown that

$$\tilde{\delta}(-1,-1) = \tilde{\delta}(0,0) = \tilde{\delta}(1,1) = \frac{1}{2}\tilde{\delta}(1,0) + \frac{1}{2}\tilde{\delta}(0,1). \tag{A.8}$$

By definition, $\Delta r + \Delta a \in [-2,2]$. Let $-2 < \Delta r + \Delta a < -1$. By aggregability,

$$\tilde{\delta}(\Delta r,\Delta a) = (2+\Delta r+\Delta a)\tilde{\delta}\left(-\frac{1+\Delta a}{2+\Delta r+\Delta a}, -\frac{1+\Delta r}{2+\Delta r+\Delta a}\right) - (1+\Delta r+\Delta a)\tilde{\delta}(-1,-1) \tag{A.9}$$

and

$$\tilde{\delta}\left(-\frac{1+\Delta a}{2+\Delta r+\Delta a}, -\frac{1+\Delta r}{2+\Delta r+\Delta a}\right) = \frac{1+\Delta a}{2+\Delta r+\Delta a}\tilde{\delta}(-1,0) + \frac{1+\Delta r}{2+\Delta r+\Delta a}\tilde{\delta}(0,-1), \tag{A.10}$$

[20] Here we assume that $\Delta r \in [-1,1]$.

which together give

$$\tilde{\delta}(\Delta r, \Delta a) = (1+\Delta a)\tilde{\delta}(-1,0) + (1+\Delta r)\tilde{\delta}(0,-1) - (1+\Delta r+\Delta a)\tilde{\delta}(-1,-1). \tag{A.11}$$

Using (A.7) and (A.11), we have

$$\tilde{\delta}(\Delta r, \Delta a) = \tilde{\delta}(0,0) + \frac{\tilde{\delta}(0,-1) - \tilde{\delta}(-1,0)}{2}(\Delta r - \Delta a). \tag{A.12}$$

Now let $-1 \leq \Delta r + \Delta a < 0$. By aggregability,

$$\tilde{\delta}(\Delta r, \Delta a) = -(\Delta r + \Delta a)\tilde{\delta}\left(-\frac{\Delta r}{\Delta r + \Delta a}, -\frac{\Delta a}{\Delta r + \Delta a}\right) + (1 + \Delta r + \Delta a)\tilde{\delta}(0,0) \tag{A.13}$$

and

$$\tilde{\delta}\left(-\frac{\Delta r}{\Delta r + \Delta a}, -\frac{\Delta a}{\Delta r + \Delta a}\right) = \frac{\Delta r}{\Delta r + \Delta a}\tilde{\delta}(-1,0) + \frac{\Delta a}{\Delta r + \Delta a}\tilde{\delta}(0,-1), \tag{A.14}$$

which together give

$$\tilde{\delta}(\Delta r, \Delta a) = -\Delta r\tilde{\delta}(-1,0) - \Delta a\tilde{\delta}(0,-1) + (1 + \Delta r + \Delta a)\tilde{\delta}(0,0). \tag{A.15}$$

Using (A.7) and (A.15), we have

$$\tilde{\delta}(\Delta r, \Delta a) = \tilde{\delta}(0,0) + \frac{\tilde{\delta}(0,-1) - \tilde{\delta}(-1,0)}{2}(\Delta r - \Delta a), \tag{A.16}$$

as in (A.12).

Now let $0 < \Delta r + \Delta a \leq 1$. By aggregability,

$$\tilde{\delta}(\Delta r, \Delta a) = (\Delta r + \Delta a)\tilde{\delta}\left(\frac{\Delta r}{\Delta r + \Delta a}, \frac{\Delta a}{\Delta r + \Delta a}\right) + (1 - \Delta r - \Delta a)\tilde{\delta}(0,0) \tag{A.17}$$

and

$$\tilde{\delta}\left(\frac{\Delta r}{\Delta r + \Delta a}, \frac{\Delta a}{\Delta r + \Delta a}\right) = \frac{\Delta r}{\Delta r + \Delta a}\tilde{\delta}(1,0) + \frac{\Delta a}{\Delta r + \Delta a}\tilde{\delta}(0,1), \tag{A.18}$$

which together give

$$\tilde{\delta}(\Delta r, \Delta a) = \Delta r\tilde{\delta}(1,0) + \Delta a\tilde{\delta}(0,1) + (1 - \Delta r - \Delta a)\tilde{\delta}(0,0). \tag{A.19}$$

Using (A.8) and (A.19), we have

$$\tilde{\delta}(\Delta r, \Delta a) = \tilde{\delta}(0,0) + \frac{\tilde{\delta}(1,0) - \tilde{\delta}(0,1)}{2}(\Delta r - \Delta a). \tag{A.20}$$

Now consider the case where $1 < \Delta r + \Delta a < 2$. By aggregability,

$$\tilde{\delta}(\Delta r, \Delta a) = (2 - \Delta r - \Delta a)\tilde{\delta}\left(\frac{1-\Delta a}{2 - \Delta r - \Delta a}, \frac{1-\Delta r}{2 - \Delta r - \Delta a}\right) + (\Delta r + \Delta a - 1)\tilde{\delta}(1,1) \tag{A.21}$$

and

$$\tilde{\delta}\left(\frac{1-\Delta a}{2 - \Delta r - \Delta a}, \frac{1-\Delta r}{2 - \Delta r - \Delta a}\right) = \frac{1-\Delta a}{2 - \Delta r - \Delta a}\tilde{\delta}(1,0) + \frac{1-\Delta r}{2 - \Delta r - \Delta a}\tilde{\delta}(0,1), \tag{A.22}$$

which together give

$$\tilde{\delta}(\Delta r, \Delta a) = (1-\Delta a)\tilde{\delta}(1,0) + (1-\Delta r)\tilde{\delta}(0,1) + (\Delta r + \Delta a - 1)\tilde{\delta}(1,1). \tag{A.23}$$

Using (A.8) and (A.23), we have

$$\tilde{\delta}(\Delta r, \Delta a) = \tilde{\delta}(0,0) + \frac{\tilde{\delta}(1,0) - \tilde{\delta}(0,1)}{2}(\Delta r - \Delta a), \tag{A.24}$$

as in (A.20).

Let $\alpha = \tilde{\delta}(-1,-1) = \tilde{\delta}(0,0) = \tilde{\delta}(1,1)$ and $\beta = (\tilde{\delta}(0,-1) - \tilde{\delta}(-1,0))/2 = (\tilde{\delta}(1,0) - \tilde{\delta}(0,1))/2$. Since $\tilde{\delta}(\Delta r, \Delta a) : [-1,1] \times [-1,1] \to \mathcal{R}$, we know that $\alpha, \beta \in \mathcal{R}$. By ordinality, it can be shown that $\beta \in \mathcal{R}_{++}$. Finally, note that this construction is consistent in the remaining cases where $\Delta r + \Delta a \in \{-2, 0, 2\}$.

Q.E.D.

A.5. *Nested Bounds on Demand Responses*

Since a nested structure implies nested empirical implications, any bounds on demand responses are also nested. Given a set of observations (p_t, q_t), $t = 1, \ldots, T$, at hypothetical prices p_0 and expenditure x_0, the set of q-rationalizable goods bundles S^q is defined according to

$$S^q = \{q_0 : U_s \leq U_t + \lambda_t p_t'(q_s - q_t),\ U_t \in \mathcal{R},\ \lambda_t \in \mathcal{R}_{++}\ \forall\ s,t = 0, \ldots, T\}, \tag{A.25}$$

and the set of z-rationalizable goods bundles S^z is defined according to

$$S^z = \{q_0 : V_s \leq V_t + \lambda_t \pi_t'(A'q_s - A'q_t),\ p_t \geq A\pi_t,\ V_t \in \mathcal{R},\ \lambda_t \in \mathcal{R}_{++},\ \pi_t \in \mathcal{R}^J\ \forall\ s,t = 0, \ldots, T\}. \tag{A.26}$$

Note that $p_t \geq A\pi_t$ is binding for any $q_{kt} > 0$. Since $\pi_t' A' \leq p_t'$, it is easy to see that the conditions in (A.26) imply that $V_s \leq V_t + \lambda_t p_t'(q_s - q_t)\ \forall\ s,t = 0, \ldots, T$, such that $S^z \subseteq S^q$.

REFERENCES

Afriat, S. N. (1967): "The Construction of Utility Functions from Expenditure Data," *International Economic Review*, 8, 67–77.

Andreoni, J., and W. T. Harbaugh (2008): "Power Indices for Revealed Preference Tests," unpublished paper.

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

Beatty, T., and I. Crawford (2009): "How Demanding Is the Revealed Preference Approach to Demand?," unpublished paper; *American Economic Review* (forthcoming).

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

Blow, L., M. Browning, and I. Crawford (2008): "Revealed Preference Analysis of Characteristics Models," *Review of Economic Studies*, 75, 371–389.

Blundell, R., and J.-M. Robin (2000): "Latent Separability: Grouping Goods Without Weak Separability," *Econometrica*, 68, 53–84.

Blundell, R. W., M. Browning, and I. A. Crawford (2003): "Nonparametric Engel Curves and Revealed Preference," *Econometrica*, 71, 205–240.

Blundell, R., M. Browning, and I. Crawford (2007): "Improving Revealed Preference Bounds on Demand Responses," *International Economic Review*, 48, 1227–1244.

Blundell, R., M. Browning, and I. Crawford (2008): "Best Nonparametric Bounds on Demand Responses," *Econometrica*, 76, 1227–1262.

Bronars, S. G. (1987): "The Power of Nonparametric Tests of Preference Maximization," *Econometrica*, 55, 693–698.

Browning, M., and P. A. Chiappori (1998): "Efficient Intra-Household Allocations: A General Characterization and Empirical Tests," *Econometrica*, 66, 1241–1278.

Chiappori, P.-A. (1988): "Rational Household Labor Supply," *Econometrica*, 56, 63–89.

Chiappori, P.-A. (1992): "Collective Labor Supply and Welfare," *Journal of Political Economy*, 100, 437–467.

Deaton, A., and J. Muellbauer (1980): "An Almost Ideal Demand System," *American Economic Review*, 70, 312–326.

Diewert, W. E. (1973): "Afriat and Revealed Preference Theory," *Review of Economic Studies*, 40, 419–425.

Fostel, A., H. E. Scarf, and M. J. Todd (2004): "Two New Proofs of Afriat's Theorem," *Economic Theory*, 24, 211–219.

Gorman, W. M. (1959): "Separable Utility and Aggregation," *Econometrica*, 27, 469–481.

Gorman, W. M. (1980): "A Possible Procedure for Analysing Quality Differentials in the Egg Market," *Review of Economic Studies*, 47, 843–856.

Griffith, R., and Nesheim, L. (2010): "Estimating Households' Willingness to Pay," unpublished paper.

Griffith, R., and M. O'Connell (2009): "The Use of Scanner Data for Research into Nutrition," *Fiscal Studies*, 30, 339–365.

Houthakker, H. S. (1950): "Revealed Preference and the Utility Function," *Economica*, 17, 159–174.

Lancaster, K. J. (1966): "A New Approach to Consumer Theory," *Journal of Political Economy*, 74, 132–157.

Leicester, A., and Z. Oldfield (2009): "Using Scanner Technology to Collect Expenditure Data," *Fiscal Studies*, 30, 309–337.

Leontief, W. (1947a): "A Note on the Interrelation of Subsets of Independent Variables of a Continuous Function with Continuous First Derivatives," *Bulletin of the American Mathematical Society*, 53, 343–350.

Leontief, W. (1947b): "Introduction to a Theory of the Internal Structure of Functional Relations," *Econometrica*, 15, 361–373.

McFadden, D. (1973): "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers of Econometrics*, ed. by P. Zarembka. New York: Academic Press, 105–142.

Polisson, M., and J. Quah (2011): "A Note on Discreteness and Revealed Preference," unpublished paper.

Richter, M. K. (1966): "Revealed Preference Theory," *Economica*, 34, 635–645.

Samuelson, P. A. (1938): "A Note on the Pure Theory of Consumer's Behaviour," *Economica*, 5, 61–71.

Samuelson, P. A. (1948): "Consumption Theory in Terms of Revealed Preference," *Economica*, 15, 243–253.

Selten, R. (1991): "Properties of a Measure of Predictive Success," *Mathematical Social Sciences*, 21, 153–167.

Strotz, R. H. (1957): "The Empirical Implications of a Utility Tree," *Econometrica*, 25, 269–280.

Varian, H. R. (1982): "The Nonparametric Approach to Demand Analysis," *Econometrica*, 50, 945–973.

Varian, H. R. (1983): "Non-Parametric Tests of Consumer Behaviour," *Review of Economic Studies*, 50, 99–110.

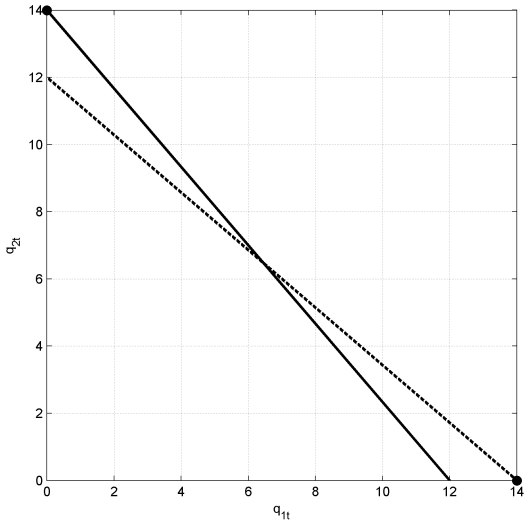


Figure 1: Example of a q-Rationalization

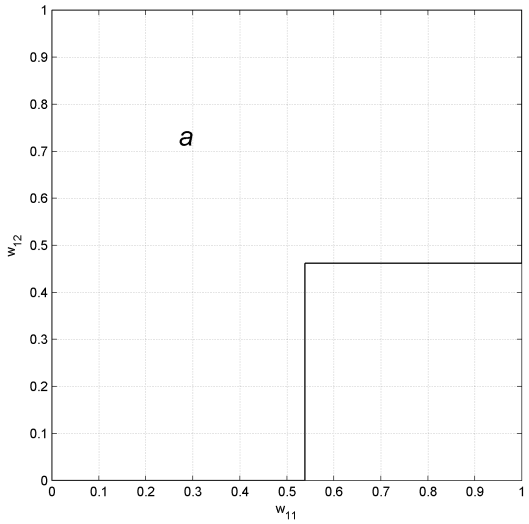


Figure 2: Example of an Area

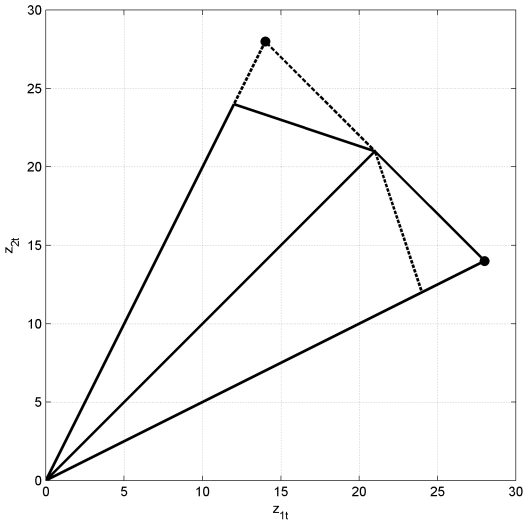
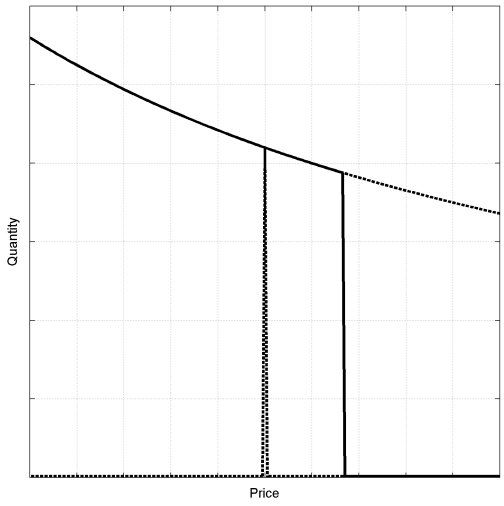
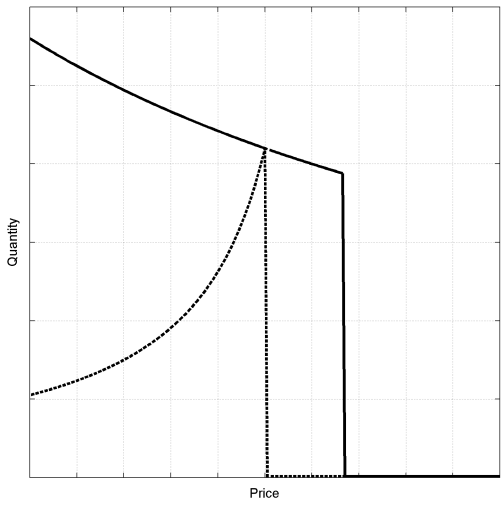


Figure 3: Example of a z-Rationalization

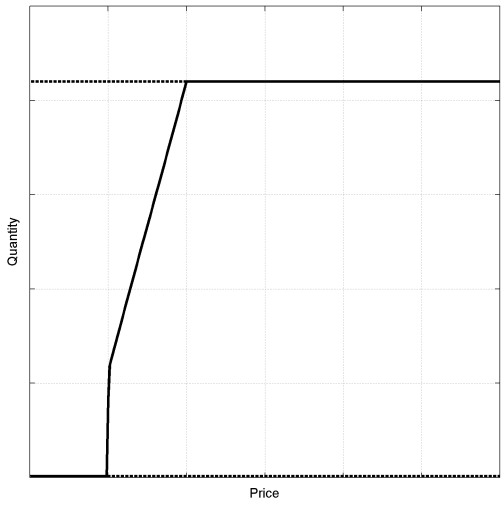


(a) Goods

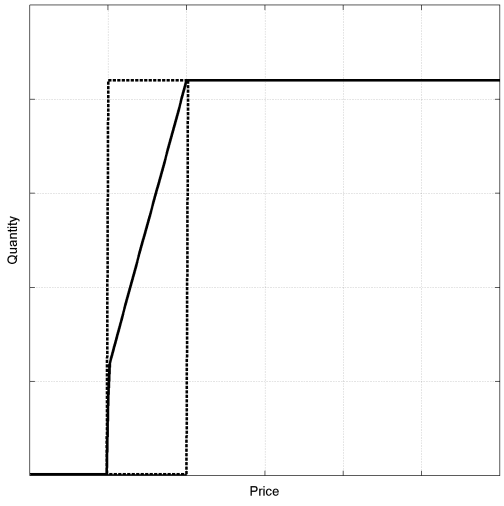


(b) Characteristics

Figure 4: Simulated Own-Price Demand Bounds

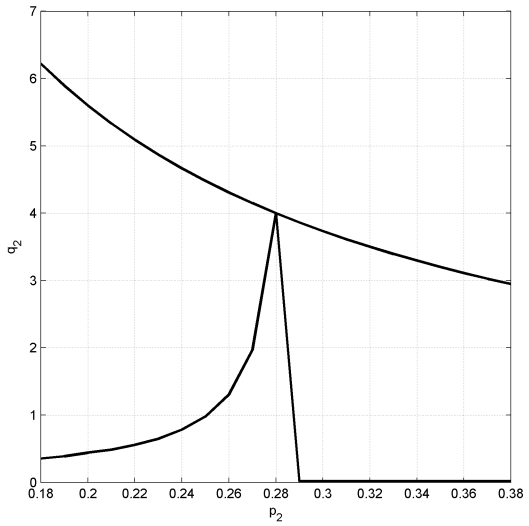


(a) Goods

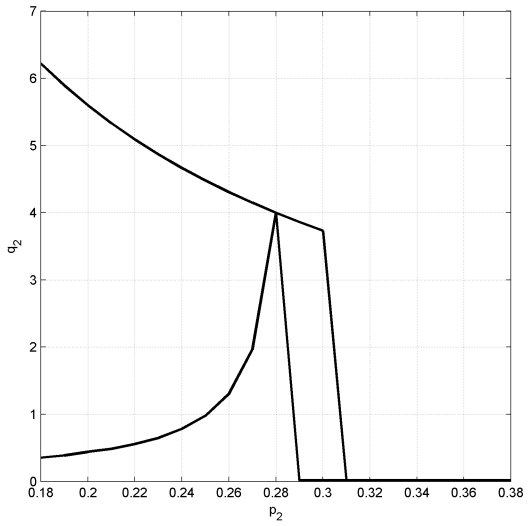


(b) Characteristics

Figure 5: Simulated Cross-Price Demand Bounds

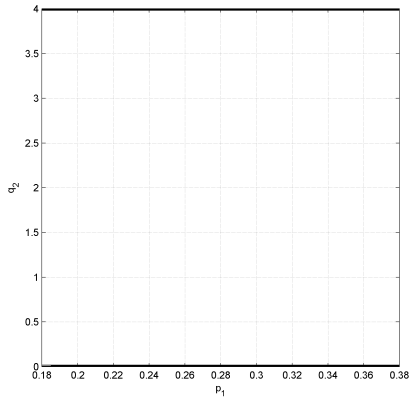


(a) Goods

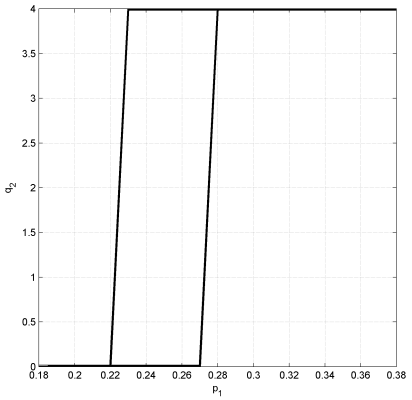


(b) Characteristics

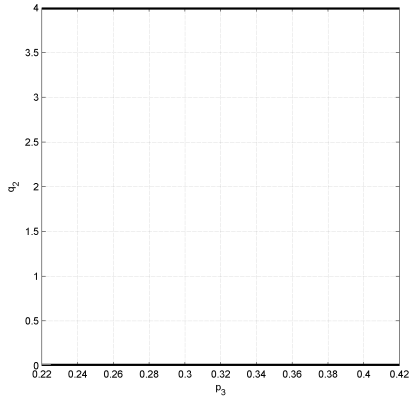
Figure 6: Own-Price Demand Bounds



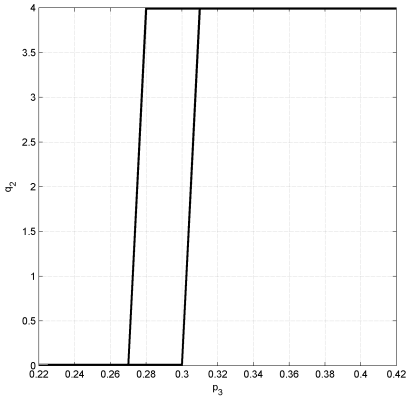
(a) Goods



(b) Characteristics

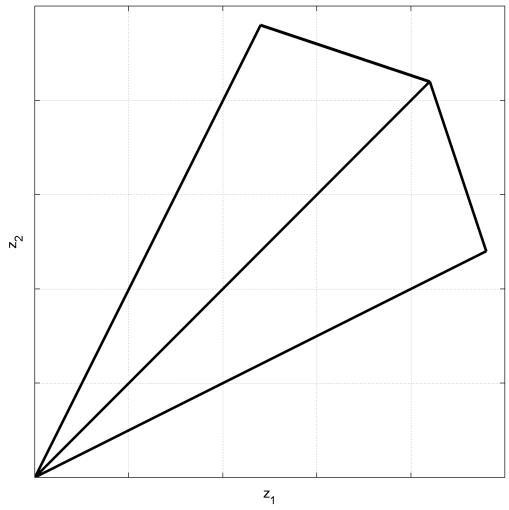


(c) Goods

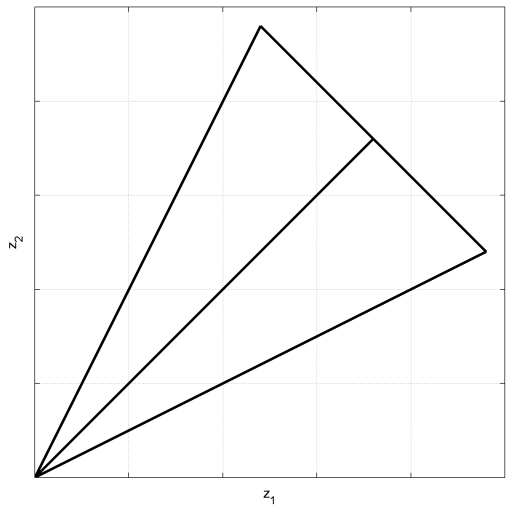


(d) Characteristics

Figure 7: Cross-Price Demand Bounds



(a) Restricted



(b) Unrestricted

Figure 8: Examples of Characteristics Budget Sets

TABLE I

MARKET SHARES

Market Share (%)	Whole	Semi-Skimmed	Skimmed
By Volume	35.2	56.8	8.0
By Value	35.0	56.7	8.3

TABLE II

PRICES

Price (£/pint)	Whole	Semi-Skimmed	Skimmed
Mean	0.31	0.31	0.32
Median	0.29	0.29	0.32
Std	0.04	0.04	0.04

TABLE III

CONSISTENCY RESULTS

	Imputed Prices		Missing Prices	
	q-rationalization	z-rationalization	q-rationalization	z-rationalization
Pass	5,084	2,935	5,516	4,987
	(90.0%)	(52.0%)	(97.7%)	(88.3%)
Fail	564	2,713	132	661
	(10.0%)	(48.0%)	(2.3%)	(11.7%)

TABLE IV

POWER RESULTS

Priors	r^q	r^z	a^q	a^z	$m(r^q, a^q)$	$m(r^z, a^z)$	$\delta(\Delta r, \Delta a)$
P1	0.900	0.520	0.857	0.000	0.043	0.520	0.477
P2	0.977	0.883	0.948	0.861	0.029	0.022	−0.007
P3	0.900	0.520	0.743	0.134	0.157	0.386	0.229