Sparse demand systems: corners and complements

Arthur Lewbel

Boston College and IFS

and

Lars Nesheim

CeMMAP, IFS and UCL

November 2019

Abstract

We propose a demand model where consumers simultaneously choose a few different goods from a large menu of available goods, and choose how much to consume of each good. The model nests multinomial discrete choice and continuous demand systems as special cases. Goods can be substitutes or complements. Random coefficients are employed to capture the wide variation in the composition of consumption baskets. Non-negativity constraints produce corners that account for different consumers purchasing different numbers of types of goods. We show semiparametric identification of the model. We apply the model to the demand for fruit in the United Kingdom. We estimate the model's parameters using UK scanner data for 2008 from the Kantar World Panel. Using our parameter estimates, we estimate a matrix of demand elasticities for 27 categories of fruit and analyze a range of tax and policy change scenarios.

JEL: C13, C34, D12, L40, L66

Keywords: sparse demand, discrete choice, continuous choice, complements, complementarity, substitutes, demand estimation, scanner data, fruit, quadratic utility

Correspondence: l.nesheim@ucl.ac.uk

Acknowledgement: We gratefully acknowledge financial support from the ESRC through the ESRC Centre for Microdata Methods and Practice (CeMMAP) grant number ES/I034021/1 and the European Research Council (ERC) under ERC-2009-AdG grant agreement number 249529. Data supplied by Kantar Worldpanel. The use of Kantar Worldpanel data in this work does not imply the endorsement of Kantar Worldpanel in relation to the interpretation or analysis of the data. All errors and omissions remained the responsibility of the authors.

1 Introduction

We propose a demand model that has features of both discrete multinomial choice and traditional continuous demand systems. In the model consumers simultaneously choose a small number of different goods from a large menu of available goods, and choose how much to consume of each good. The model has wide applicability in large scale demand estimation settings in which most consumers choose zero demand for most goods.

Our model nests both standard continuous demand systems (quadratic utility functions with Gorman (1976, 1980) and Lancaster (1966) consumption technologies) and standard discrete choice models (multinomial logit or probit with random coefficients) as special cases. Unlike most discrete choice models, our model allows the chosen goods to be substitutes or complements, and to be consumed in continuous quantities. Unlike standard continuous consumer demand systems, our model allows individual consumers to choose zero quantities of most types of goods, and includes substantial unobserved preference heterogeneity. A key feature of our model is that we treat the output of the Gorman-Lancaster linear consumption technology as a vector of unobserved latent indexes, allowing for highly flexible patterns of substitution or complementarity.

As our motivating example, we consider consumer demand for fresh fruit in the UK. In a typical store, there are more than two dozen types of fruit that consumers can choose among. Consumers typically choose from one to five different types of fruit to purchase, and buy varying quantities of each type. Some types of fruits are substitutes (such as apples vs bananas) while others are complements (like cantaloupe and honeydew melons in fruit salad). Some fruits might be substitutes for some households while being complements for others. The types and quantities of fruits purchased vary greatly across households.

While many different types of fruit are offered for sale, typical households only buy a small number of types. As a result, most consumers buy zero quantities of most categories of fruit, and therefore the vector of observed demands at the individual consumer level is sparse. Note that this is not a model that is sparse in the sense of having many zero coefficients,

like regressions estimated using the Tibshirani (1996) LASSO estimator. Rather, here it is the data that is sparse, since for each shopping trip, each consumer buys zero quantities of most of the goods that are available in the store.

The most popular method of dealing with such sparse demand systems, as exemplified by the Berry, Levinsohn, and Pakes (1995) BLP model, is to discretize purchases and treat each unit purchased as an independent multinomial choice decision. Unfortunately, in many empirical applications not only is this method intractable but also the standard assumptions underlying the methodology are likely to be seriously violated. For example, in our empirical application consumers buy up to 5 different types of fruits from a set of 27 available types of fruits. So even ignoring the quantities purchased and only looking at the types of fruit selected, there are 80,730 possible baskets to consider, which is far too large for traditional discrete choice methods.

A far more serious limitation of multinomial choice models is that they generally rule out complements. Complementarities are important in a wide range of empirical applications. For example, in our application, some fruits are strong complements (e.g., different types of berries are frequently purchased and consumed jointly, and various fruits are complementary inputs to dishes like fruit salad). It is possible to allow for complements in a discrete choice framework by modelling combinations of fruit as additional distinct goods, e.g., treating an apple, a banana, and the combination of both as three separate possible choices. However, the number of possible combinations of just a few fruits out of more than two dozen makes this approach impractical. We could alternatively allow for some complementaries in a reduced form way by assuming logit shocks that are correlated across purchase decisions, but the number of such correlations would again become rapidly intractable.

The leading alternative to multinomial choice models of demand for many goods are traditional continuous demand models such as those described in Deaton and Muellbauer (1980). These models are designed to handle joint purchases of bundles of goods in continuous quantities. However, such models assume each consumer buys positive quantities of most or

all goods. Methods exist for dealing with small numbers of zeros in such models (essentially, system Tobit; see, e.g., Yen and Lin 2006 and references therein). However, in our sparsity case each consumer buys zero amounts of a large majority of the available goods.

When using traditional demand systems, large numbers of zeros are usually dealt with by aggregating to form a few broad categories of goods. However, such aggregation leads to biases of unknown size and direction unless strict aggregation conditions are met. The separability or price co-movement restrictions required to justify Gorman or Hicksian aggregation (see, e.g., Lewbel (1996) and references therein) often do not hold. Moreover, for many applications in marketing, industrial organization, public finance, or in health, one is interested in the determinants of demand for each type of fruit, not just for broad aggregates. In Section 7.4 we give an example in which the introduction of a tariff on EU sourced fruits affects each category of fruit differently depending on the fraction sourced in the EU, and we compute the disparate impacts on each type of fruit.

The basic structure of our model incorporates a Gorman (1976, 1980) and Lancaster (1966) type linear consumption technology into a continuous demand system with substantial unobserved preference heterogeneity. The model then allows for many corner solutions in the demand for unobserved characteristics to account for the sparsity of observed individual consumer demands, while the heterogeneity allows different consumers to be at different corner solutions.

Our model has J different kinds of goods, and contains K latent indices that are linear functions of consumption quantities (in our fruit application, $J = 27$ and $K = 5$). As a result, K is the maximum number of types of goods that any consumer will purchase at one time (except for knife edge situations of indifference). The number of different types of goods a particular consumer actually purchases at any one time, which ranges from zero to K, is determined by the number of nonnegativity constraints that bind (i.e., the number of corners) in the consumer's utility maximization problem. When maximizing utility, the consumer simultaneously determines how many different types of goods to buy, which goods

to buy, and the quantity to purchase of each good.

In one limiting case where $K = J$, our model reduces to a standard continuous choice quadratic utility model, where all available goods are purchased in continuous quantities. At the other extreme, when $K = 1$ our model reduces to the Dubin and McFadden (1984) model where consumers choose a single good by standard multinomial choice (e.g., probit), and also choose to purchase a continuous quantity of that good. An alternative limit case of our model nests standard multinomial logit or probit models as special cases. Our model therefore nests standard multinomial choice (with or without random coefficients), standard continuous demand systems, and classic mixed continuous and discrete demand models all as special cases. As a result, our model has wide application across a range of demand estimation settings.

The next section is a literature review. Section 3 lays out our model, and Section 4 shows how our model nests standard continuous, discrete, and mixed models as special cases. Section 5 gives our semiparametric identification results, and describes our estimator. Sections 6 and 7 describe our fruit demand application and our empirical results. Section 8 concludes. A separate Supplementary Appendix provides additional technical material, summary statistics and estimation details.

2 Literature

As summarized by, e.g., Blundell and Meghir (1987), the continuous demand literature considers three main theoretical rationales to explain zero expenditure on some goods. One rationale is lexicographic preferences. With lexicographic preferences, an individual might prefer to consume any amount of other goods, no matter how small, to a given good. A second rationale is infrequency of purchase due to durability or storage. A good that is durable or storable may be consumed regularly, but infrequently purchased. In our fruit demand example, infrequency of purchase can be largely ruled out over time spans longer

than a few days, because fresh fruit is not durable and cannot be stored for very long. A third rationale is corner solutions. These occur when the price of a good is above its reservation price so that nonnegativity constraints are binding. In such cases, given prices and total expenditures, a consumer chooses to purchase zero units of the good in question.

Lexicographic preferences are typically modelled analogously to Heckman (1979) type sample selection models. A binary choice equation models the decision of whether to consume the good or not, and then ordinary demand systems are estimated either including or excluding the good in question. Systems of equations like these can be estimated parametrically using Shonkwiler and Yen (1999) or Yen and Lin (2006). A recent example (still with a small number of goods) is the semiparametric estimator of Sam and Zheng (2010). Models like these require utility functions that are fundamentally different for non-consumers and consumers of a good. These types of models are generally most appropriate for goods that a significant fraction of the population would never consume, like tobacco or alcohol.

In our model we focus on corners, since it is likely that very few types of fruit are goods that households would never purchase. In addition, our model allows for substantial preference heterogeneity, and so accommodates the types of behaviour that lexicographic preferences seek to capture by allowing some consumers to have arbitrarily small marginal utility for some goods. Our model allows for the possibility that purchases of some goods may be extremely rare for a significant fraction of households.

Extreme versions of models based on corners are brand choice models where the constraint that consumers buy exactly one brand is imposed either a priori or by the structure of the utility function. For example, Hendel (1999) proposes a model in which firms choose a single brand (of computer) along with a number of units (firms that are observed to buy multiple brands are divided into separate tasks, and each task is treated as if it was an individual firm choosing one brand). Similarly Dube (2004) proposes a model where the purchase for each "consumption occasion" is the decision to purchase a single brand, but in a continuous quantity. Other models that entail choosing a single good among many and consuming that

good in continuous quantities include Dubin and McFadden (1984) and Haneman (1984), and more recently Crawford and Yurukoglu (2012).

A drawback of all these discrete choice based models is that they rule out the possibility of many different goods being complements. None would, e.g., allow for the possibility of making a fruit salad. In contrast, our model is based directly on continuous joint demand for multiple goods, and so allows for goods to be complements, and more generally places no separability restrictions on the demands for different goods.

Corners in continuous demand models are generally modelled as censored regressions, such as Tobit models. The early continuous demand system literature that considered corners formally focused on cases where either a single good, or a very small number of goods, may have zeros. Examples include Wales and Woodland (1983) and Lee and Pitt (1986). Applications of continuous demand systems with many goods and censoring generally work as follows. Let p and y be a price vector and total expenditures, respectively. Utility maximization without nonnegativity constraints are first used to derive models of the form $q_j^* = f_j(p, y) + e_j$ for each good J, where q_j^* is a latent quantity and e_j is an error term. Each observed quantity q_j is then assumed to be given by $q_j = \max\{0, q_j^*\}$. Examples of such models include Golan, Perloff, and Shen (2001) and Meyerhoefer, Ranney, and Sahn (2005).

These censored demand models have one of two flaws. Either errors e_j are arbitrarily appended to demand functions yielding empirical specifications of the form $f_j(p, y) + e_j$, or errors are incorporated as random utility parameters but ignored in estimation. That is, demand equations of the form $q_j^* = f_j^*(p, y, e) + e_j$ are approximated by $f_j^*(P, Y) + e_j$. The most common example of this latter method is based on Deaton and Muellbauer's (1980) Almost Ideal Demand System (AIDS), where the vector e appears in the demand functions $f_j^*(P, Y, e)$ only inside a general price index as in Heien and Wessells (1990).

Most of these censored continuous demand models are not fully consistent with utility maximization because the nonnegativity constraints are not explicitly incorporated into the consumer's utility maximization. In these models, the consumer first chooses possibly nega-

tive quantities for some goods to maximize utility, and then actually purchases zero quantities for these goods. These problems apply to almost all demand systems with many goods that allow for censoring based either on e or those based on separate selection equations. An exception is the brand choice models that forbid complementarities discussed earlier, which solve this problem by imposing extreme forms of separability. Difficulties in preserving regularity in demand models with non-negativity constraints are further discussed in Van Soest, Kapteyn, and Kooreman (1993), and Millimet and Tchernis (2008).

Continuous demand models do exist where random utility parameters e are not removed by approximation (see, e.g., Lewbel and Pendakur 2009, 2017), but censored versions of these models have mostly not been developed. An exception is Amano (2018), who essentially applies Lee and Pitt's (1986) theory to Lewbel and Pendakur (2009) EASI model, employing a simulated method of moments estimator to overcome analytical difficulties. However, this approach becomes impractical when the number of goods is large. Amano (2018) must therefore still maintain strong two stage budgeting assumptions, and model at the level of aggregate categories of food, to avoid having too many categories of food containing zeros.

Two other papers that have looked at complementarities across goods are Beckert, Griffith, and Nesheim (2009) and Thomassen, Smith, Schiraldi, and Seiler (2017). The former paper develops a discrete choice store choice model in which, at the second stage, consumers choose a basket of goods, possibly including zero demand for some goods, using a quadratic utility model. Unlike our model, their's does not allow for many goods with many corners, and does not include random coefficients. Their empirical analysis aggregates goods to a high level and is limited to an application with 4 types of goods.

The latter paper, Thomassen et al. (2017), develops and estimates a store choice model allowing consumers to purchase from multiple stores and accounting for zeros in demand. The paper develops and estimates the implications of complementarities and multi-store shopping behaviour for competition analysis. As in our paper, the consumer utility model is quadratic.

Our paper adds to the frameworks developed in these previous papers by allowing the quadratic utility model to be less than full rank, allowing for flexible heterogeneity to affect both the first and second derivatives of utility, and by analysing demand at a much more disaggregate level. The additional flexibility in heterogeneity is required to match variation in baskets across households.

In our model, zeros are handled using both corners and the Gorman (1976, 1980) and Lancaster (1966) linear consumption technology with taste heterogeneity. Dubois, Griffith, and Nevo (2014) also exploit a Gorman Lancaster technology, but with observable characteristics and only to account for taste heterogeneity for types of food, and not for dimension reduction. Theirs is a continuous demand system, and so despite enormous sample sizes they must still substantially aggregate across goods to avoid zeros (e.g, they treat spending on all fruits as a single aggregate good). Other papers that allow for many corners with Gorman-Lancaster observable characteristics are Chan (2006) and Kim, Allenby, and Rossi (2007), but these models impose strong additivity conditions that rule out complements along with most other forms of interactions between goods. In addition, the restriction that characteristics are observable greatly limits the flexibility of these models. Also related (in terms of its factor structure) is Elrod and Keane (1995), though theirs is a discrete choice probit model.

The model we propose overcomes all of the problems summarized above. Each consumer takes all nonnegativity constraints directly into account when maximizing utility. The model directly incorporates error terms as preference heterogeneity parameters and allows for arbitrary patterns of substitutability or complementarity among the goods. The model allows consumers to buy continuous quantities of some goods and zero quantities of the rest. The model is broadly applicable to any situation where consumers choose multiple options from a large discrete choice set.

3 The model

Let q_j be the quantity of good j purchased by a consumer or a household, and let $q \in \mathbb{R}_+^J$ be the bundle of goods purchased by this consumer. Later we add a subscript h to index households, but for now, omit that to simplify notation. Suppose that consumer utility from q is a function of K latent attributes. Let b_{kj} be the quantity of attribute k that a consumer derives from buying a unit of good j and let B be the $K \times J$ matrix of elements b_{kj}. Then the K vector of attributes a consumer derives utility from is the vector Bq. Assume $K \leq J$, $\text{rank}(B) = K$ and $B^T B \geq 0$. This is essentially the Gorman-Lancaster linear household technologies model.

We assume consumers have a strictly quasiconcave utility function over the K dimensional latent attributes Bq. The particular functional form we use for this utility function is quadratic. The quadratic utility assumption is not necessary for the analysis but offers numerical simplicity when applied to large scale datasets.

For now, we assume all consumers have the same matrix B. Later, we introduce observable (demographics) and unobservable (random coefficient) heterogeneity into B. This heterogeneity will be important empirically to capture the fact that consumers facing the same prices choose different baskets of goods.

In a standard continuous demand model, each consumer generally buys nonzero quantities of all J goods. However, in the Gorman Lancaster model, utility is maximized by consumers buying exactly K different types of goods. One feature of our model is that we let K be much smaller than J, which then accounts for most of the zeros in our data. A second feature is that we introduce preference variation across consumers in the form of random terms that are added to each element of the vector of latent attributes Bq (later, we also introduce additional variation in the form of random coefficients). This preference variation across consumers results in different consumers choosing different baskets of goods. Even with this taste heterogeneity, the Gorman model would be inadequate for real data, because it implies that each consumer, with probability one, buys the same number of different types

of goods, K.

An additional feature of our model is that we allow that maximized utility may have many corners, i.e., points where indifference curves intersect axes in attribute space. As a result, depending on prices and preference parameters, utility may be maximized by choosing anywhere from zero to K different types of goods. The more corners (the more binding constraints), the smaller is the optimal number of different goods to purchase.

Analogous to a Tobit model, in our model the marginal value of each latent index (i.e., the marginal utility from each element of Bq) plus unobserved heterogeneity determines whether a given attribute is desired sufficiently (relative to its cost) to purchase in nonzero amounts. The unobserved heterogeneity terms are location shifts in the marginal utility for each attribute. The interaction of these preference heterogeneity terms with binding corners results not only in different consumers purchasing different baskets of goods, but also in different consumers facing different corners, and different numbers of corners. The result is that our model can encompass the variation seen in data, where consumers vary in the numbers of goods that they buy (from zero to K), vary in the choice of which goods to buy, and vary in the quantities they purchase of each nonzero good.

Assume that each individual chooses q to maximize the utility function

$$u\left(q_0\right) - 0.5\left(e - Bq\right)^T\left(e - Bq\right) \text{ such that } y \geq p^T q + q_0 \text{ and } q \geq 0$$

where u is a monotonically increasing function, $y \in \mathbb{R}_+$ is total grocery expenditures, $q_0 \in \mathbb{R}$ is a numeraire good, $p \in \mathbb{R}_+^J$ is a price vector, and $e \in \mathbb{R}^K$, which is randomly distributed in the population, is a vector of preference parameters, where element e_k corresponds to a satiation level or bliss point for attribute k.

This utility function is quadratic and weakly concave in q which allows us to employ standard efficient quadratic programming techniques to handle zeros coming from corner solutions. These methods are computationally fast even for very large quadratic programs.

Importantly for large scale estimation, it also allows us to efficiently analyse the inverse of demand and compute the probabilities of observing the data at hand as functions of model parameters. The theory would largely go through with more general utility functions that are concave in $e - Bq$, but would be computationally more burdensome.

This utility function nests both standard continuous demand systems and standard discrete choice models. We discuss this equivalence in more detail in Section 4. We also discuss incorporating additional observable and unobservable heterogeneity in a way that includes random coefficients multinomial probit or logit as special cases.

For the rest of the paper we let $u(q_0) = q_0$, making preferences quasilinear and thereby eliminating income effects. This simplification is reasonable for our empirical application, since fruit and vegetables are generally a small component of households' overall budgets. Assuming quasilinear utility, normalizing the marginal utility of income to be one[1], and substituting the budget constraint into the objective, the consumer chooses q to maximize

$$y - p^T q - 0.5 (e - Bq)^T (e - Bq) \text{ such that } q \geq 0. \tag{3.1}$$

3.1 First order conditions

The Lagrangian for each consumer's maximization problem is

$$L(q, \delta) = y - p^T q - 0.5 (Bq - e)^T (Bq - e) + \delta^T q$$

where δ is a vector of Lagrange multipliers. The first order conditions are

$$\begin{aligned} 0 &= -p - B^T (Bq - e) + \delta \\ 0 &= \delta^T q, \ \ \delta \geq 0, \ \ q \geq 0. \end{aligned} \tag{3.2}$$

[1]The utility function in (3.1) can be multiplied by any positive number without changing any predictions or implications of the model.

By assumption, the second order conditions are satisfied since $-B^T B \leq 0$.

Due to quasilinearity, the value of y does not affect the optimal choice of q. This model implicitly assumes either that the numeraire can be consumed in negative quantities, or that $y \geq p^T q$ for any optimizing value of q. Note that this latter condition holds automatically as long as y is large enough to purchase a bundle q that attains the satiation level $Bq = e$ (though consumers in that situation may still not choose to buy that bundle, if the utility value of holding more of the numeraire is greater).

This model has the property that any consumer can maximize utility by buying nonzero amounts of at most K goods. Given prices and B, the first order conditions define a partition of $\mathbb{R}^K$ with at most $R = \binom{J}{K}$ elements and where each element of the partition is a polytope. Let E_r be an element of this partition. All consumers with $e \in E_r$ choose a quantity q_r with the same non-zero components. For each consumer, calculating their optimal quantity bundle q entails solving a concave quadratic program. Finding an optimum requires identifying the relevant element of the partition and then computing the optimal quantity. Because the problem is a concave quadratic program, algorithms exist that obtain a solution in polynomial time (interior point and related methods). For estimation, computing the likelihood function requires finding the set E_r corresponding to each demand observation q and then computing the probability that $e \in E_r$. Because E_r is a polytope, we are able to construct efficient algorithms to compute this probability. Details are in the Supplementary Appendix.

3.2 Piecewise linear demands

To prove identification in Section 5.1, it is useful to characterize solutions that have the maximum number K of nonzero elements. To do so, let $\overline{q} = (\overline{q}_1, \overline{q}_2)$ be a vector for which $\overline{q}_1 > 0$ and $\overline{q}_2 = 0$ such that $\dim(\overline{q}_1) = K$. Without loss of generality, the elements of $\overline{q}_1$ can be taken to be the first K elements of $\overline{q}$. Let p_1 and p_2 be the corresponding price subvectors

and B_1 and B_2 the corresponding submatrices of B so that

$$B = \left[\begin{array}{cc} B_1 & B_2 \end{array} \right].$$

That is, B_1 is the $K \times K$ matrix formed from the first K columns of B and B_2 is the $K \times J - K$ matrix formed from the remaining $J - K$ columns.

Then $\overline{q}$ is optimal for all e satisfying

$$\begin{aligned} -p_1 - B_1^T \left(B_1 \overline{q}_1 - e \right) &= 0 && (3.3) \\ -p_2 - B_2^T \left(B_1 \overline{q}_1 - e \right) &\leq 0 && (3.4) \\ \overline{q}_1 &\geq 0 && (3.5) \end{aligned}$$

Equation (3.3) defines the inverse demand curve $-p_1 = B_1^T \left(B_1 \overline{q}_1 - e \right)$ for a single consumer. The inequalities (3.4) and (3.5) define conditions under which choosing $\overline{q}_1 > 0$ and $\overline{q}_2 = 0$ is optimal. When B_1 is nonsingular the system can be simplified and solved to provide explicit conditions describing the piecewise linear demand function

$$\begin{aligned} \overline{q}_1 &= \left(B_1^T B_1 \right)^{-1} \left(B_1^T e - p_1 \right) && (3.6) \\ p_2 - B_2^T \left(B_1^T \right)^{-1} p_1 &\geq 0 && (3.7) \\ \overline{q}_1 &\geq 0. && (3.8) \end{aligned}$$

In words, the nonnegative $\overline{q}_1$ is optimal if it satisfies the demand equation (3.6) and if the projection of its price vector is cheaper than the price vector p_2. When (3.4) is not binding, small changes in p_2 have no impact on demand for $\overline{q}_1$.

4 Special Cases

In this section we show that our model nests standard continuous choice, discrete-continuous choice and discrete choice demand models as special cases. In particular, different types of continuous and multinomial choice demand systems result from setting $K = J$, $K = 1$, or by imposing limiting constraints on B.

4.1 Continuous consumer demand

Suppose $K = J$. Then the model simplifies to an ordinary continuous quasilinear quadratic utility function, which (by the first order conditions derived earlier) yields the demand equations

$$q = \left(B^T B\right)^{-1} \left(B^T e - p + \delta\right), \; 0 = \delta^T q, \;\; \delta \geq 0, \, q \geq 0.$$

where δ are Lagrange multipliers. When all elements of $B^T e - p$ are nonnegative then the nonnegativity constraints do not bind and so with $K = J$, the system of linear continuous demand equations given by (3.6) becomes

$$q = \left(B^T B\right)^{-1} \left(B^T e - p\right).$$

For empirical application, one could then let B or e depend on product characteristics z, consumer characteristics x, or unobserved heterogeneity η as detailed in Section 5.2 below. For example, one could assume $e_{jh} = (\beta_0 + \beta_1 x_h) z_j + \varepsilon_{jh}$ to obtain a linear demand system over continuously demanded goods.

4.2 Discrete-continuous choice

Before considering standard discrete choice, it is useful to examine how our model relates to Dubin and McFadden (1984). They propose a model in which each consumer chooses a single type of good according to a multinomial probit model and then purchases a continuous

quantity of the chosen good. Suppose that $K = 1$ in our model. Then B equals the row vector of nonnegative elements b_{1j}, e equals the scalar e_1, and the consumer's problem of equation (3.1) reduces to

$$\max \left\{ y - p^T q - 0.5 \left(Bq - e_1 \right)^2 \right\} \tag{4.1}$$

Since $K = 1$, utility is maximized by purchasing at most one type of good. The consumer's utility from buying q_j units of good j is

$$y - p_j q_j - 0.5 \left(b_{1j} q_j - e_1 \right)^2 ,$$

which is maximized either at an interior point of the feasible range of values of q_j given by the first order condition $-p_j - \left(b_{1j} q_j - e_1 \right) b_{1j} = 0$ or at one of the endpoints of the feasible range, i.e., either $q_j = 0$ or $q_j = y/p_j$. At an interior solution, the optimal quantity is $q_j = \left(e_1 - p_j / b_{1j} \right) / b_{1j}$ (which is only feasible if $e_1 > p_j / b_{1j}$) which yields utility $y + 0.5 \left(\left(p_j / b_{1j} \right) - e_1 \right)^2 - 0.5 e_1^2$. Otherwise, at the corner solution, if $q = 0$ is optimal, then utility is $y - 0.5 e_1^2$.

It follows that if $e_1 \leq \min_{\ell \in \{1, \ldots, J\}} \{ p_\ell / b_{1\ell} \}$ for all goods j, then utility is maximized by $q = 0$. Otherwise, utility is maximized by buying the quantity $q_j = \left(e_1 - p_j / b_{1j} \right) / b_{1j}$ of the good $j = \arg\min_{\ell \in \{1, \ldots, J\}} \{ p_\ell / b_{1\ell} \}$, and not buying any other good. Let $j = 0$ denote not buying any of the goods. Let $p_0 = 1$ and $\ln b_{10} = -\ln e_1$. It then follows that equation (4.1) implies making a discrete choice to purchase good $j = \arg\max_{\ell \in \{0, 1, \ldots, J\}} \{ \ln b_{1\ell} - \ln p_\ell \}$ and a continuous choice of q_j as detailed above.

As discussed in Section 5.2 below, in empirical applications the elements of B may depend on consumer and product characteristics. For example, one could assume $\ln b_{1jh} = \left(\beta_0 + \beta_1 x_h \right) z_j + \varepsilon_{jh}$ where z_j and x_h are vectors of observed product and consumer characteristics. Letting $\varepsilon_{0h} = \ln e_{1h}$, our model reduces to a multinomial choice model in which a

consumer chooses to purchase only the good j that satisfies

$$j = \arg \max_{\ell \in \{0,1,\ldots,J\}} \{\beta_h z_\ell - \ln p_\ell + \varepsilon_{\ell h}\}.$$

(which is multinomial probit if ε is normal) and the quantity q_j given by $q_{jh} = (e_{1h} - p_j/b_{1jh})/b_{1jh}$.

4.3 Multinomial discrete choice

The previous section shows one way our model encompasses discrete-continuous choice. Additionally, a limiting case of our model nests ordinary multinomial choice or pure discrete choice. Suppose we take B to be a diagonal, invertible $J \times J$ matrix. Consider the model where $q \geq 0$ is determined by maximizing the utility function

$$\left(y - p^T q\right) \beta_0 - 0.5 q^T B^T B q + u^T q \tag{4.2}$$

for some positive scalar β_0 (which equals the marginal utility of money) and vector of fixed or random parameters u. When B is invertible, this model is equivalent to our model, because for any choice of B and u, one can define $e = B^{-1}u$, which then makes equation (4.2) equal to equation (3.1) up to an affine transformation (multiplication by β_0 and addition of $e^T e$) that has no effect on consumer choices. Essentially, our original utility model can be derived from equation (4.2) by completing the square.

Now consider the limiting case of (4.2) in which $B \to 0$ and $Be \to u$ where $u_j = \beta z_j + \varepsilon_j$. In the limit case, maximizing (4.2) is equivalent to maximizing

$$\sum\nolimits_{j=1}^{J} \left(\beta z_j - \beta_0 p_j + \varepsilon_j\right) q_j.$$

With the addition of the constraint that $\sum_j q_j \leq 1$, this is the standard multinomial choice model. The parameters (β, β_0), may depend on both observable demographics x_h and random coefficients η_h. In particular, if one assumes that ε_j are distributed as independent

Type I extreme value random variables, then this is the standard multinomial logit model.

5 Identification and estimation

In Section 5.1 we maintain the assumption that all consumers have the same B and discuss conditions sufficient to ensure that B is point identified. We also consider nonparametric identification of the distribution of the vector e. We show that this distribution is point identified for all e in the set $\left\{e : B^T e \geq 0\right\}$. For values not in this set, the distribution is not identified because consumers with values of e outside this set, choose $q = 0$ with probability one. To see this, consider a consumer with e satisfying $B^T e \leq 0$. For this consumer, it follows immediately from the first order conditions (3.2) that $q = 0$ for all nonnegative prices. Therefore, if prices are nonnegative, nothing can be identified regarding the distribution of e for all $e \in \{e \mid B^T e \leq 0\}$, other than the probability of lying in this set.

In Sections 5.2 and 5.3 we discuss heterogeneity in B before discussing estimation in Section 5.4.

5.1 Identification

ASSUMPTION A1: *With probability one, consumers buy the minimum number of different goods necessary to maximize utility given by equation (3.1). Assume that p is continuously distributed on the positive orthant with a density that is strictly positive almost everywhere on the positive orthant. Assume that the distribution of q given p is known.*

The assumption that consumers buy the minimum number of goods is a tie breaker for knife edge situations where utility can be maximized in more than one way. Given the assumed continuity of prices, these knife edges occur with probability zero. The distribution of q in a population facing prices p is in principle observable, so Assumption A1 essentially says that, for proving identification, this distribution is assumed to be known for any value of p.

ASSUMPTION A2: *The $K \times J$ matrix B has rank $K > 0$. For every column B_j of B, there exists a $(K \times K-1)$ matrix B_{-j} consisting of $K-1$ columns of B such that $\widetilde{B}_j = \left[\begin{array}{cc} B_j & B_{-j} \end{array} \right]$ is nonsingular. Without loss of generality, B is assumed to be upper triangular.*

Assumption A2 ensures that for every good j, there exists a set of K goods including good j such that some consumers choose to buy a bundle consisting of those K goods. Identification of the j'th column of B is assured using expressions like (3.6) and (3.7) with nonsingularity of $\widetilde{B}_j$ in Assumption A2 taking the place of nonsingularity of B_1.

For any $K \times K$ matrix A such that $A^T A = I$, our utility function is observationally equivalent to a utility function that replaces B and e with AB and Ae. Specifically, B can only be identified up to a set of scale and rotation normalizations. That is, the scale (or magnitude) of each column of B is identified as is an upper triangular matrix defining, within each column, the relative magnitudes of the elements of B. These normalizations can be imposed by assuming B is upper triangular.

ASSUMPTION A3: *Let f_e denote the density function of e. The density f_e is strictly positive on the set $E = \left\{ e \mid B^T e \geq 0 \right\}$. e is distributed independently of p.*

The price exogeneity Assumption A3 is a strong restriction. We discuss this along with price measurement issues later in a "Price issues" subsection.

THEOREM 1: Given Assumptions A1, A2, and A3, the density $f_e(e)$ is nonparametrically identified for all $e \in E$ and the matrix B is identified.

Proof of Theorem 1: Let $\mathcal{B}$ be the set of unique combinations of K different goods chosen from the J available goods, and let $R = \binom{J}{K}$ be the total number of elements of $\mathcal{B}$. Let $r \in \{1, ...R\}$ index each possible element of $\mathcal{B}$. Let $i^r = \{i_1^r, ..., i_k^r\}$be an element of $\mathcal{B}$ and let q_r be a vector of quantities satisfying $q_r(i) = 0$ if $i \notin i^r$. We call q_r a K dimensional basket or bundle corresponding to the list i^r. So for a given basket q_r, i^r indexes the nonzero elements

of q_r. Let p_r be the K vector of prices of the goods q_r, and let p_{-r} be the $J-K$ vector of prices of all the other goods. Let $B^r = B(:, i^r)$ be the submatrix of B corresponding to these nonzero elements. Let $\widetilde{R} \subset \mathcal{B}$ denote the smallest set of bundles such that B^r is nonsingular for all $r \in \widetilde{R}$ and B_j is a column in B^r for some r. The set $\widetilde{R}$ has at least J/K elements and no more than $J-K+1$ elements. By Assumption A2, for every good j the column B_j lies in some nonsingular B^r.

With these definitions, we first show that for every $r \in \widetilde{R}$, there is a set $A \subseteq P \times Y$ and a set $Q^r = \left\{q_r \in \mathbb{R}^J \text{ with } q_r(i) = 0 \text{ if } i \notin i^r\right\}$ such that $\Pr(Q^r | A) > 0$. To show this, consider $q_r \in Q^r$. It is optimal to choose q_r when inequalities (3.7) and (3.6) are satisfied for $\overline{q} = q_r$. That is when $p_{-r} - B_{-r}^T \left(B_r^T\right)^{-1} p_r \geq 0$ and $q_r = \left(B_r^T B_r\right)^{-1} \left(B_r^T e - p_r\right) \geq 0$. Assumptions A2 and A3 ensure that this event has positive probability.

Given this result, we can now establish identification of B. For each good j, there is a subset $\mathcal{B}_r$ of K goods as described above that includes good j. For this set of goods let p_r be sufficiently low, and let p_{-r} be sufficiently high, to yield a positive probability of observing bundles q^r in which $q^r(i) > 0$ for all $i \in \mathcal{B}_r$. Then $q^r > 0$ for all $p' = \left(p'_r, p'_{-r}\right)$ where $p'_r \leq p_r$ and $p'_{-r} \geq p_{-r}$ (p', y).

Let B_r be the $K \times K$ submatrix consisting of the columns of B corresponding to the set $\mathcal{B}_r$ of these K goods, and let p_r and $\overline{q}_r$ denote K vectors of prices and quantities of those K goods. By the first order conditions, a consumer buying $\overline{q}_r$ has $B_r^T B_r \overline{q}_r = B_r^T e - p_r$. By assumption A2, $B_r^T B_r$ is nonsingular. The demand functions for these K goods for the consumers in this region are therefore $\overline{q}_r = \left(B_r^T B_r\right)^{-1} \left(B_r^T e - p_r\right)$. Since the distribution of e does not depend on p_r, the derivative with respect to prices p_r of the conditional mean (or any conditional quantile) of $\overline{q}_r$ conditioning on p (which can be calculated at any point that is not on the boundary of the region) is $\left(B_r^T B_r\right)^{-1}$, which identifies $B_r^T B_r$.

By Assumption A2, each good j appears in some bundle r for which the above derivation can be performed and $B_r^T B_r$ can be identified, so all of the columns of B are recoverable up to normalizations from the collection of estimates of $B_r^T B_r$. At most $J-K$ such bundles r

would be required (so that each good j appears in at least one such bundle) and as few as J/K such bundles might be needed.

We have now shown that for each r, we can identify

$$A_r = B_r^T B_r.$$

In addition, these matrices share common elements. So, we can pick one bundle r and define

$$A_r = D_r C_r C_r^T D_r$$

where D_r is a positive diagonal matrix and C_r is the Cholesky decomposition of a correlation matrix. We can then define

$$B_r = C_r^T D_r.$$

This provides the rotation and scale normalizations up to which B is identified. Given $B_r = C_r^T D_r$, the remaining columns of B are identified by sequentially dropping the last column of B_r and replacing it with each remaining column of B. The elements of column j for $j \notin \mathcal{B}_r$ satisfy $\sum_i [B_j(i)]^2 = d_j^2$ for some $d_j > 0$.

Having now shown identification of B, consider the distribution of e. Given B_r for all possible bundles r, we can observe $B_r^T e = B_r^T B_r \overline{q}_r + p_r$ for all observable $\overline{q}_r, p_r$ pairs. Since $\overline{q}_r$ and p_r are nonnegative, we can uncover observations of $B_r^T e$ and hence of e for all $e \in E$, thereby identifying $f_e(e)$ for all $e \in E$. QED.

Theorem 1 shows that $f_e(e)$ is identified for $e \in E$. As discussed earlier, for $e \notin E$, it is not possible to learn anything about $f_e(e)$ other than the total probability of not lying in the set E. The people with values of $e \notin E$ are never willing to pay a positive price to buy fruit.

For policy questions such as competition policy questions or tax policy questions, these people are irrelevant. They are outside the market. For policy questions involving exter-

nalities such as public health, they may be relevant and policy makers may be interested in learning about the distribution of e for those consumer types. In that case, policy makers have several options. They could estimate bounds on policy responses, they could introduce experiments with subsidies to generate negative prices, or they could identify the distribution by imposing shape or parameter restrictions on f_e. Since $f_e(e)$ is identified over a large (positive measure) subset of the support of e, in general it could be fully identified either by semiparametric shape restrictions such as radial symmetry, or by finitely parameterizing the density. It then follows by Theorem 1 that the model is completely identified.

5.2 Heterogeneity in B

We now introduce heterogeneity in B and discuss how B shapes product choices and the degree of substitutability and complementarity between goods. The main reason we introduce heterogeneity in B is empirical - there appears to be more variation in consumption patterns, holding prices fixed, than can be explained by variation in e (analogous to the use of random coefficients in addition to logit errors in applications of BLP models).

Let B^h be the matrix of preference parameters for household h, let B^h_j be column j in B^h and let b_{hkj} be row k column j in B^h. As discussed in Section 4.2, when $K = 1$, the ratio of p_j to $|b_{h1j}|$ determines product choice for household h. Household h purchases the good with the smallest value of $\frac{p_j}{|b_{h1j}|}$. In this case, all goods are perfect substitutes and goods with low prices and large values of $|b_{h1j}|$ are purchased.

When $K > 1$, the magnitude of each column vector $\|B^h_j\| = \sqrt{\sum_{k=1}^{K} (b_{hkj})^2}$, plays a similar role. When $\|B^h_j\|$ is large relative to p_j, the product j is likely to be purchased. However, now households may buy more than one good and goods may be complements. Now, the relative magnitudes of the elements within a column of B^h, determine how important each good is in producing each latent attribute. They also govern the degree to which goods are complements or substitutes.

As discussed above, B^h is identified only up to scale and rotation normalizations.[2] To impose these normalizations while incorporating heterogeneity in a flexible way, we parameterize B^h as follows. First, we normalise B^h to be upper triangular so that $b_{hkj} = 0$ if $k > j$. Then we convert the nonzero elements of each column of B^h into hyperspherical coordinates. That is, for each j, we define $\left(d_{hj}, C_j^h\right) = H\left(B_j^h\right)$ where $d_{hj} = \|B_j^h\|$, $\|C_j^h\| = 1$, and H is the hyperspherical coordinate transformation detailed in Appendix A in the Supplementary Appendix. In the hyperspherical coordinate representation $d_{hj} \in \mathbb{R}_+$, $c_{hkj} \in [0, \pi]$ for all $k \leq \min(K, j) - 2$, and $c_{hkj} \in [0, 2\pi]$ for $k = \min(K, j) - 1$. Here, c_{hkj} is the element in row k of vector C_j^h. Finally, we assume that

$$\begin{aligned} \ln d_{hj} &= z_j^T \beta_h && (5.1) \\ c_{hkj} &= \pi \Phi^{-1}\left(z_{kj}^T \gamma_h\right) \ \forall k \leq \min(K, j) - 2 && (5.2) \\ c_{hkj} &= 2\pi \Phi^{-1}\left(z_{kj}^T \gamma_h\right) \ \forall k = \min(K, j) - 1 && (5.3) \end{aligned}$$

where Φ is the normal CDF, (z_j, z_{kj}) are vectors of product characteristics, and (β_h, γ_h) are vectors of consumer specific parameters. The log transformation ensures that d_{hj} is positive and the inverse normal transformations ensure that c_{hkj} are constrained to lie in the relevant intervals.[3]

We assume

$$\begin{aligned} \beta_{hj} &= \beta_{j0} + \beta_{j1}^T x_h + \beta_{j2}^T \eta_h && (5.4) \\ \gamma_{hkj} &= \gamma_{kj0} + \gamma_{kj1} x_h + \gamma_{kj2} \eta_h && (5.5) \end{aligned}$$

where x_h is a vector of observable demographic variables and η_h is a N_η dimensional vector of unobservable latent factors normalized to be independent with mean zero and variance 1.

[2]The presence of heterogeneity means that these normalization are applied separately for each household. If the distribution of random coefficients were nonparametric, these would be free normalizations, however we do impose a functional form on the distribution, though our model of heterogeneity is flexible.

[3]The inverse normal transformation could be replaced with the inverse of any strictly increasing cumulative distribution function.

The $J \times N_\eta$ matrix $\beta_2 = \left[\beta_{12}, ..., \beta_{j2}, ..., \beta_{J2}\right]$ is an upper triangular matrix of factor loadings mapping the low dimensional η into the random coefficients β_h. Let $N_C = (K-1)\left(J - \frac{K}{2}\right)$ be the number of elements of $\{C_j\}_{j=2}^J$ Then, the $N_C \times N_\eta$ matrix $\gamma_2 = \left[\gamma_{122}, ..., \gamma_{KJ-1,2}\right]$ is an upper triangular matrix of factor loadings mapping the low dimensional η into the random coefficients c_h. Note that the matrices of factor loadings are normalized to be upper triangular and that the parameters describing the mean, variance and correlations of η are subsumed in the parameters $(\beta_0, \beta_2, \gamma_0, \gamma_2)$.

In terms of product substitutability, this specification nests the fully unrestricted case in which (z_j, z_{kj}) are vectors of dummy variables defined by product names and more restricted case where (z_j, z_{kj}) are vectors of observable product characteristics. The former case is unrestricted in the sense that no patterns of substitutability are imposed on B^h. The latter, depending on the set of observable characteristics available, imposes that products with similar values of (z_j, z_{kj}) are similar.

In terms of heterogeneity across households, the flexibility of the model depends on the set of observable variables available and on the dimension and distribution of η. In our empirical application we assume that the dimension of η_h is 2 and that it is normally distributed.

The model is highly flexible in that the support of the random coefficients spans the space of upper triangular matrices B. In addition, we allow random coefficients to affect both the importance of each product $\|B_j\|$ and the patterns of substitution and complementarity. This flexibility is important to capture the wide variety of baskets chosen by households. Finally, we maintain this flexibility while keeping the dimension of random coefficients low by using the factor structure in equations (5.4) and (5.5).

5.3 Identification of additional heterogeneity

The proof of Theorem 1 works by establishing nonparametric identification of the distribution of $B^T e$ in the positive orthant. When B is constant, Theorem 1 shows this implies identification of B and nonparametric identification of the distribution of e for all e where

$B^T e > 0$. These results remain true if both B and f_e are conditioned on x_h.

In our empirical application, we also introduce random coefficients into B as detailed in the previous section. To maintain identification, we assume that the distributions of both the random coefficients and e are finitely parameterized, with distributions that are independent of p.

The proof of Theorem 1 shows that for various subsets r of K goods, people who purchase positive amounts of those K goods (because the price vector p_r of those goods is sufficiently low and the price vector of all other goods is sufficiently high) do so with demand functions given by $\overline{q}_r = \left(B_r^T B_r\right)^{-1}\left(B_r^T e - p_r\right)$. Conditioning on this price regime, the conditional distribution of $\overline{q}_r$ given prices nonparametrically identifies the distribution of $\left(B_r^T B_r\right)^{-1}\left(B_r^T e - p_r\right)$ given p_r.

The parameterized distributions of B and e are then identified as long as their parameters can be recovered from moments of $\left(B_r^T B_r\right)^{-1}$ and of $\left(B_r^T B_r\right)^{-1} B_r^T e$. In our empirical application, we assume e is a multivariate K vector normal and, as detailed in the previous section, we introduce random coefficients into B using a factor structure. Specifically, we assume η_h is a N_η dimensional vector of independent normally distributed latent factors. We then estimate the factor loadings β_{j2} and γ_{kj2}. The identification of the parameters of normal distributions from low order moments then ensures identification of this parameterized model.

5.4 Estimation

We assume the data consists of n independent observations of (p_{ht}, q_{ht}, x_{ht}) for each household (h,t). Observation (h,t) is an observation of household h in month t.[4] Income y_{ht} plays no role due to the quasilinear utility assumption. However, income can be included as a household characteristic, as is common in empirical applications of multinomial demand

[4]We use one observation per household and do not exploit the panel structure of the dataset. While it is straightforward conceptually to extend the analysis to the panel setting under standard assumptions, the computational burden increases and so we leave the analysis to future work.

models. To capture seasonal patterns in demand, we assume that $e \sim N(\mu_t, \Sigma)$ with $\mu = \{\mu_t\}_{t=1}^T$ and that equation (5.4) is replaced with

$$\beta_{hjt} = \beta_{j0t} + \beta_{j1}^T x_h + \beta_{2j}^T \eta_h \tag{5.6}$$

β_{j0t} varies across months. In short, we allow aggregate demand for each of the K latent indexes to vary across months and we allow the relative importance of each good, β_{j0t} to vary across months. We assume that the parameters γ do not vary across months. Finally, we assume that $\eta \sim N(0, I)$, and that $\dim(\eta) = 2$. The assumption that η has mean zero and covariance matrix equal to the identity matrix is a normalization given the parameterization in equations (5.4) and (5.5).

The parameters of the model are $\theta = (\mu, \Sigma, \beta, \gamma)$ where $\mu = \{\mu_t\}_{t=1}^T$ is the vector of all mean values of e, $\beta = \left\{\left\{\beta_{j0t}\right\}_{t=1}^T, \beta_{j1}, \beta_{j2}\right\}_{j=1}^J$ is the vector of all parameters in (5.1) and $\gamma = \left\{\gamma_{kj0}, \gamma_{kj1}, \gamma_{kj2}\right\}$ is the vector of all parameters in (5.2). The parameters (μ, Σ) determine the distribution of e. They primarily determine the number of items chosen and the quantities purchased. The parameters (β, γ) determine the distribution of B. They govern which products are chosen and in which combinations. They also determine the response to prices. We estimate the model parameters by maximum likelihood.

Full details of the log likelihood function are given in Appendix B in the Supplementary Appendix. Here, we simply outline the key elements. We compute the likelihood function in each of three cases. Case 1 is the case where a consumer purchases exactly K goods. In this case, conditional on the random coefficients, the mapping from data to e is one-to-one. The likelihood function is simply that of a linear model with random coefficients and computation merely requires integration with respect to the distribution of random coefficients. Because we assume a factor structure on the random coefficients, the dimension of integration is kept low.

Case 2 is the case where a consumer chooses fewer than K items but more than zero.

In this case, conditional on random coefficients, many values of e are consistent with the observed choice and so the likelihood function is the integral over the polytope in $\mathbb{R}^K$ defined by the first order conditions. To compute the integral efficiently, we make a series of change of variables to convert the integral to an integral over a hypercube and then use the tensor product of Gauss-Legendre integration rules to compute the integral on the hypercube. Because the region of integration is a polytope in the original coordinates, the change of variables is simple and fast to compute. A benefit of the change of variables is that the boundary of the transformed region of integration does not depend on the parameters so the numerical approximation preserves the fact that the likelihood function is a smooth function of the parameters.

Case 3 is the case where a consumer chooses to purchase nothing. This case is similar to case 2 but with a slightly different set of binding inequalities defining the region of integration.

6 Empirical application

We use data from the Kantar World Panel for the UK for calendar year 2008 on all purchases of food brought into the home by 26,514 households. Using handheld scanners, households record purchases of all items bought and record prices from till receipts. We treat each shopping trip as an observation. The data contain a large set of product attributes (at the barcode level) as well as household characteristics. We use data on all purchases of fruit excluding a few infrequently purchased categories. After eliminating these small categories, we observe purchases of 27 different types of fruit including, for example, apricots, bananas, apples, and cherries.

6.1 Summary statistics

Table A.1 shows the purchase frequency of each category of fruit. The top three most frequently purchased categories are bananas (23.79 % of purchases), apples (16.85%) and

grapes (9.99%). The top 15 categories account for 95% of purchases.

Table A.2 shows the purchase frequency of different sized baskets. The table shows that 48.18% of baskets contained exactly 1 item (that is, any quantity of one type of fruit), 25.63% contained two items and 13.86% contained 3 items. Households purchased baskets containing 5 or fewer items 97.67% of the time. A simple discrete choice model that assumes consumers buy at most one type of fruit, would be wrong 51.82% of the time.

Table A.3 shows the most frequently purchased two-item combinations. While each of the top 5 or 10 two-item combinations has an appreciable market share, in aggregate the top 5 account for only 54.34% of two-item combinations and the top 10 account for only 67.20%. To account for 95% of two-item combinations one must include 105 distinct combinations, which are all the combinations listed in Tables C.1-C.3 in the Supplementary Appendix. Most of these combinations have small market shares individually, but together they account for a large share of all two-item baskets. Our model can account for this wide variation in choices of types of fruit, numbers of types chosen, and the quantities of each.

Another way to see the variety of choices and the potential role of complementarities is to look at the frequency of basket size conditional on fruit choice. Tables C.4-C.5 show, conditional on purchase of a fruit type, how frequently each basket size was purchased. Except for bananas, cherries, and lemons, all categories are more likely to be purchased in combinations than as stand-alone categories. The relative frequencies of basket size vary across fruit categories and the larger baskets are usually less frequent. These patterns strongly violate the usual independence assumptions of typical discrete choice demand models.

6.2 Prices

For every shopping trip, we observe the expenditure and price for all items purchased. However, for items not purchased the price is not observed. To overcome this problem, we follow standard practice and impute prices using a hedonic regression. For each fruit category we

estimate a hedonic price model

$$\ln p_{it} = \beta x_{it} + h(t) + \varepsilon_{it}$$

where $\ln p_{it}$ is the price of item i in period t, x_{it} is a vector of characteristics of item i in period t and $h(t)$ is a 6th order polynomial function of time. Time is measured as the day within the year. Characteristics included in the regressions are country of origin, branded, organic, tiering (economy, premium or standard), fascia (one of ten firms in the UK or other), packaging, online shop, and small store.

Figure D.1 in the Supplementary Appendix shows price data and imputed prices for 3 representative examples of the 27 fruit categories: apricots, bananas and cherries. Price is observed for each shopping trip where a particular fruit is purchased. For apricots and cherries, prices rise in the spring and the autumn. These are periods when fresh apricots and cherries are more costly and more scarce. In contrast, the price of bananas is relatively flat. The results also show that, at any single point in time there is a great deal of variability in price. This variation is primarily due to variation across fascia and variation due to promotions.

6.3 Price issues

Issues regarding prices include imputations, seasonal unavailability, and potential endogeneity. For some stores and time periods, no purchases of a particular fruit are observed. As noted above, we follow standard procedure in the literature by imputing prices for these periods using a hedonic pricing model.

For some time periods, the number of observed purchases is either zero or extremely low. Fruits (such as ugli fruit) that have very low demand in all time periods are dropped from the sample, since demand for these is too low and specialized to be estimated with reasonable precision. Other fruits have very low or zero sales just in some time periods but not others,

due to availability (typically being out of season). For example, cherries are available in the summer and winter but not in the spring and the autumn. To handle this situation, we trim observations when sales within a one week window are below a low threshold. That is, if total sales of a fruit within a week in our sample are below the threshold, we treat the fruit as unavailable that week and drop from our sample the few households that did manage to purchase that week. For these cases, we treat the fruit's price as being arbitrarily high on those days, to represent lack of availability.

Essentially, this procedure treats low availability as a supply shock. We interpret this procedure as a form of asymptotic trimming. By lowering the threshold as the sample size grows, asymptotically we only treat true zeros as unavailable supply, noting that any infrequency of purchase or high demand price elasticity will eventually lead to some purchases in every period where the product is really generally available.

The estimation of our model assumes prices are exogenous. Since we estimate the model using data on daily purchases, likely sources of endogeneity for fruit demand on any particular day could include promotional activity, weather (if both prices and demand respond to weather), and unobserved quality variation. Most of the variation in quality of fruit is either across stores or seasonal. We capture seasonal variation using monthly dummies in the model. We treat store choice as exogenous (noting that store choice depends heavily on factors other than fruit demand, such as distance to the store, and on the other products consumers consider buying on each trip to the market). We are therefore assuming that, conditional on store, prices are not correlated with demand shocks. Conditional on seasonal dummies, we expect current weather to shift demand but not price as we expect stores to rarely if ever change prices in response to changes in high frequency (such as daily) weather shocks. We include promotional status in our hedonic price models and assume that conditional on price, unobserved demand shocks are independent of promotional activity. In summary, given the nature of our data and the controls we include in the model, we expect that biases from assuming prices are exogenous are likely to be small.

6.4 Potential estimation issues

Our identification proof assumes that prices are continuously distributed over a relatively large support, which ensures that, with positive probability, most possible combinations of $K = 5$ or fewer fruits would be purchased by some subset of consumers. However, in finite data sets, we may not observe many combinations of less popular fruits being purchased, or the number of consumers observed buying rare combinations of fruits may be very small. An analogous problem arises in BLP type models, where some goods may have very small or zero market shares. In practice, our estimator converges and appears to be numerically well behaved, as we discuss in the next section. This may be aided in part by our use of a parametric model for the distribution of random utility parameters, which should allow for identification even with potentially limited price variation.

7 Empirical results

The total number of parameters in the model is determined by the number of types of fruit J, the number of indices K (which equals the maximum number of different types of fruit a single consumer may buy), N_η, the dimension of latent factors, and the number of month fixed effects for μ_t and β_{0jt}. For $J = 27$, $K = 5$, $N_\eta = 2$, and $T = 12$, there are 740 parameters. We estimate these parameters by maximum likelihood, using our sample of 26,514 observations. At the optimum we find that the Hessian of the likelihood function is negative definite (largest eigenvalue is -0.15) and that all parameters are estimated with a high degree of precision. All are statistically distinct from zero at the 5% level. We restarted the estimation procedure from multiple starting points, and tested by perturbing the model in multiple directions in the parameter space. We found no evidence of multiple local optimizers or of failure to converge to a global optimum.

Individual parameter values are difficult to interpret. So, to illustrate our results, we discuss model predictions for individual households with different random coefficient values,

we summarize aggregate demand curves and elasticities implied by the estimates, and we provide several different counterfactual simulation exercises.

7.1 Household level demand predictions

To illustrate model predictons for individual consumers, we plot predicted demand for 9 household types. Each type is defined by a realization of the vector η such that for each type each element of η takes one of three values. For each element, the three values are selected from the 25th, 50th or 75th percentile of the respective marginal distributions. Each type also is set to have values of e equal to the mean.

For each household type, we computed the frequency with which baskets of various sizes were purchased as prices vary (one at a time) from 50% to 200% of baseline prices (basket size here refers to the number of different types of fruit, not the quantities of each). Household types 2, 5, 7 and 8 virtually always buy exactly two types of fruit. There are only a small number of settings in which they buy 1, 3, 4 or 5 items. At the same time, household types 1 and 4 buy 2 or 3 items most of the time (they buy 1 or 4 items on a small number of occasions) and household types 3, 6, and 9 usually buy 1 or two items (they buy 2-5 items on a small number of occasions).

For each household type, we also examined how the basket composition varies with price. Household types 1-6, always buy bananas and apples but vary in terms of which fruits are added to their basket when they buy more than two items. Household type 1, 4, and 5 purchase kiwis and nectarines. Households 2, 3, and 6 buy easy-peelers, kiwis, and nectarines. Households 7-9 do not buy bananas but do buy apples, kiwis, nectarines, and easy-peelers.

Finally, Figures B.1 and B.2 plot these households' demand curves for various fruits, as functions of the prices of bananas and apples, respectively. Turning first to Figure B.1, the figures shows demand for bananas, apples and kiwis as function of the price of bananas for households 1-6. Households 7-9 are not shown because they buy no bananas; as a result the

banana price does not affect their demand. For households 1-6, the banana demand curves are downward sloping, with slope varying across households. In addition, household 3 stops buying bananas when the price rises above 1.13, and for household 4, the banana demand curve is kinked because the household starts buying kiwis when the banana price rises above 1.35. The effects of the banana price on demand for other fruits vary widely. Some cross-price effects are negative, some are positive, and some are flat. Figure B.2 shows a similar wide range of effects of apple prices on fruit demand by household type. Apple demand curves slope downward. Some individual demand curves are kinked when a household either starts or stops purchasing a type of fruit. cross-price effects are positive in some cases, negative in others, and flat when apple demand is zero.

These 9 household types illustrate the types of individual behaviour predicted by the model. However, they only illustrate a handful of cases. To further analyse the model's predictions, we next analyse aggregate demand.

7.2 Aggregate demand predictions

Figures B.3 - B.5 show estimated aggregate demand curves for each fruit. They also show what fraction of aggregate demand comes from purchases of baskets with 1 to 5 items. While the demand curves for individual consumers are piecewise linear, the variation across households in slopes and kink points produces aggregate demand curves that are smooth and show varying degrees of curvature.

The aggregate banana demand curve is shown in Figure B.3 panel (c). It has a relatively gentle and constant slope . Most of the demand for bananas comes from shoppers purchasing 3, 4, or 5 items. In contrast, panel (b) shows the avocado demand curve. The avocado demand curve is mush steeper and has much more curvature. Very few consumers who buy avocados buy 5 items. Most buy 2-4 items. These curves illustrate that both intensive and extensive margin effects influence the shapes of the aggregate demand curves.

7.3 Elasticities

Tables A.4 - A.7 show estimates of average own- and cross-price elasticities. Six of the own-price elasticities are less than one in magnitude (apples, bananas, easy-peelers, lemons, pears, and plums). Of these apples, bananas and easy-peelers are in the top 5 fruit categories in terms of market share, pears and plums are in the top 10 and lemons are 11th. This suggests that these products might at least sometimes be sold as loss leaders, on sale for a relatively low price, despite inelastic demand. Twenty of the elasticities are between -1.04 (mangos) and -10.6 (sharon fruit) and seventeen elasticities are between -1 and -5.

One of the fruit own-price elasticities is very large in magnitude; pomegranates (-41.1).This is not altogether unexpected. Pomegranates are purchased in only 0.16% of transactions. While they are not the smallest market share product, it is a feature of sparse demand heterogeneous consumers that demand for products with small market shares can have very high elasticities, because it only requires a small number of consumers to start buying the product to produce a very large percentage increase in demand. The aggregate demand curve for pomegranates seen in Figure B.5 panel (g) shows that the demand curve for pomegranates is very steep when prices are low and flattens as the pomegranates price rises. As a result, the estimated own-price elasticity drops substantially when the price rises.

Typical discrete choice models assume all goods are substitutes and so do not permit zero or negative cross-price effects. Likewise, typical continuous demand models do not have exactly zero cross-price effects. In contrast, in our model estimated cross-price effects between two types of fruit will be exactly zero (consistent with economic theory) when the two types of fruit are never purchased in the same basket. Tables A.4 - A.7 show that about a third of all pairs of fruits have zero cross-price effects at baseline prices.

Among the nonzero cross-price elasticities there are a mix of negative and positive effects. The negative cross-price elasticities indicate that on average in our sample, the goods are complements. For example, looking at row 3 in Tables A.4-A.7, when the price of bananas rises, demand for 13 of the other groups goes down (avocados, berries, cherries, dates, apples,

easy-peelers, grapes, grapefruit, kiwi, lime, mango, melon, and pineapple). At baseline prices, these goods in aggregate are complements to bananas. Demand for 10 other goods goes up (lemons, lychees, nectarines, oranges, passion fruits, paw paws, peaches, pears, plums, and pomegranates). At baseline prices, these goods in aggregate are substitutes to bananas. For some of the small market share goods, these cross-price effects are large. For example, the impact on pomegranates is 5.88. Because most people who buy pomegranates also buy bananas, the banana price has a big impact on demand for pomegranates. In contrast, the cross-price effect of pomegranate price on bananas is only 0.000482. It is small because most people who buy bananas do not buy pomegranates.

7.4 Counterfactual scenarios

Many current large scale shifts in the economy could affect the markets for fruit in the UK. For example, Brexit is likely to increase tariffs on fruit imports from Europe. Brexit could also increase the costs of UK fruit by limiting the supply of farm workers and driving up wages. Another potential change would be a merger between two of the largest supermarket chains. A proposed merger between Asda and Sainsbury's, who account for 16.8% and 15.5% of UK fruit sales respectively, was blocked by the UK competition authority in April 2019. Such a merger could increase their market power, possibly driving up fruit prices.

At the same time, various tax policy changes could be considered by the British government. Currently, due to concerns about tax incidence on poor households, purchased food to be eaten at home is not subject to the VAT (value added tax). Extending the VAT to food could significantly increase tax revenue at the cost of adversely affecting poor households. Alternatively, the government might consider subsidising fruit consumption to promote public health (in the past the British government promoted fruit consumption in other ways, such as the "five a day" advertising campaign).

To analyse effects of these potential changes, we simulated the impacts of five different policy scenarios:

1. A 10% increase in the prices of EU sourced fruit due to Brexit.

2. A 10% increase in the prices of UK sourced fruit due to Brexit.

3. A 5% increase in the price of fruit at Asda and Sainsbury's.

4. A 10% subsidy of fruit prices to promote public health.

5. A 20% VAT tax on fruit to raise revenue.

To simulate the first three scenarios, we used our sample to compute the fraction of each fruit category sourced from the EU, from the UK and from the rest of the world. We also computed the fraction of each category sold by Asda and Sainsbury's. We then used these shares to compute the price changes implied by each of these events.

For scenario one, the percentage price increase for fruit j is assumed to be $\tau_{1j} = 0.1 s_{EU,j}$ where $s_{EU.j}$ is the share of fruit sourced from the EU. For scenario two, the percentage price increase for category j is assumed to be $\tau_{2j} = 0.1 s_{UK,j}$ where $s_{UK,j}$ is the share of fruit sourced from the UK. For scenario 3, the percentage price increase is assumed to be $\tau_{3j} = 0.05 \left(s_{ASDA,j} + s_{SAIN,j} \right)$ where $s_{ASDA,j}$ and $s_{SAIN,j}$ are the shares of fruit sold by Asda and Sainsbury's respectively. For scenario four, we assume all fruit prices decrease by 10%. For scenario 5, we assume all prices increase by 20%.

The price changes resulting from each of these scenarios are detailed in Table A.8. The first two scenarios affect prices in complex ways because the fraction of fruits sourced from each country varies significantly across fruit types. For example, the EU tariff scenario results in more than a 5% price increase for apricots, kiwis, lemons, nectarines, peaches and pears, because relatively large fractions of those fruits are sourced from the EU. In contrast, the UK cost shock results in price increases of less than 5% for all fruits except rhubarb (9.87% increase). A small number of categories (berries, cherries, apples, pears,) have more moderate price increases of greater than 1%. These are the only categories for which the UK is a significant supplier. The merger has a more balanced impact on prices because there isn't much variation in fruit market shares across grocery firms.

While the exact price impacts of Brexit and of the proposed merger are unknown (see for example Levell et al. 2017), the hypothesised price changes we consider here provide an illustration of what the first order impacts from Brexit or the proposed merger could be.

For each scenario, given the change in prices, we use our model to compute the impact on a) demand, b) welfare, c) revenues of grocery firms, and d) tax revenue. Results are given in Tables A.9 and A.10.

The second column in Table A.9 shows that the EU tariff has a small percentage impact on most fruit categories but a big negative impact on apricots, cherries, nectarines, peaches and pomegranates. For all of these categories, the tariff leads to a drop in demand of more than 5%. The impacts are quite large, larger than one would predict based on the own-price elasticities alone. In addition, due to substitution effects, several categories (bananas, dates, apples, grapefruits, lychees, mangos, melons, oranges, passion fruits, paw paws, and pineapples) experience a net increase in demand. These fruits are primarily sourced either from the UK or from outside the EU. As a result, their prices are unaffected by the tariff and yet their demand increases, in some cases by substantial amounts. Taken together, these results illustrate that cross-price effects are quite important for understanding the impact of tariff shocks on demand for fruit.

Scenarios 2 and 3 have much more moderate impacts on prices and also on the resulting demand for fruit, for all fruits except rhubarb. Rhubarb demand, low to begin with, is reduced to zero by the predicted 9.87% price increase. Demand for berries, cherries, pears, and plums is reduced by -1.12% (pears) to -7.61% (berries) while demand for lemons, nectarines, pineapples, peaches, pomegranates, and sharon fruits is increased by 1.09% (sharon fruits) to 15.6% (pineapple). All other responses are less than 1%. In scenario 3, in which prices increase by 1.08% (dates) to 2.21% (apricots), most products experience changes of demand of less than 5%. The exceptions are apricots and sharon fruits which experience declines in demand of 13% and 7.37% respectively.

The final two scenarios, a 10% subsidy of fruit prices and a 20% VAT on fruit, have

large impacts on demand. The former increases demand by less than 10% for 15 categories, 10-20% for 6 categories and more than 20% for 6 categories. The latter scenario reduces demand by less than 10% for 10 categories, 10-20% for 7 categories and more than 20% for 10 categories.

Table A.10 reports impacts on total consumer expenditure and on welfare. The top panel shows the impact on household fruit expenditure per shopping trip. The first 3 scenarios lead to increases in expenditure ranging from 0.85% to 1.95%. Importantly, the change in expenditure induced by the price change is not monotonically related to total expenditure. In 4 of 5 scenarios the 90th percentile of expenditure changes the most and in 3 of 5 scenarios the 10th percentile changes the least. In scenario 1 the 50th percentile is impacted more than the 25th, whereas in scenario 2, the 25th is impacted more than the 50th. This illustrates that accounting for unobserved heterogeneity is important for capturing how price changes affect households at different points in the expenditure distribution.

The second panel shows the impact on consumer surplus measured in GBP per household per shopping trip. The EU tariff costs 10th percentile households about 5 pence per shopping trip and costs the 90th percentile households about 44 pence per shopping trip. The UK cost shock has very small impacts, less than 10 pence per shopping. The merger has an intermediate impact. Scenarios 4 and 5 lead to larger price changes and hence larger impacts on consumer surplus.

The final panel summarizes the aggregate impacts of each scenario. In all cases, the consumer surplus effects and tax revenue effects offset each other almost exactly. However, the price increases lead to reductions in firm revenue. The EU tariff reduces firm revenue by 13 pence per household per shopping trip, representing about a 2% decrease in revenues. The merger actually reduces firm fruit revenues by about 7 pence per trip and reduces consumer surplus by 11 pence per trip. Since fruit accounts for only a small share of supermarket revenue, this suggests that the merger would have increased revenue from other goods enough to compensate for this reduction in revenue.

8 Conclusions

Discrete choice models of demand focus on the fact that consumers must make individual selections from a wide variety of items in the market. However, many goods are not purchased and consumed in isolation, but jointly with other goods. Also, many goods are purchased and consumed in close to continuous quantities rather than in single units. Unlike most discrete choice models, our model allows consumers to choose more than one good at a time, allows the chosen goods to be substitutes or complements, and lets goods be consumed in continuous quantities. Unlike standard continuous consumer demand systems, our model allows individual consumers to choose to consume zero quantities of most types of goods, and includes substantial unobserved preference heterogeneity in the form of random coefficients. Our model nests both standard continuous demand systems like the quadratic direct utility function and standard discrete choice models like random coefficients logit or probit as special cases.

In our empirical application to fruit demand in the UK, we uncover a wide range of demand patterns, including complementarities, kinks, and corners, that could not have been revealed with traditional discrete or continuous demand models. These results have important implications for welfare calculations, construction of price indices, market structure, and tax policies. We illustrate some of these implications by estimating the impacts of potential policies such as tariffs or price changes due to Brexit, a change in the VAT, or a merger between two large grocery chains.

9 References

Amano, N. (2018), "Nutrition Inequality: The Role of Prices, Income, and Preferences," Yale University unpublished manuscript.

Beckert, W., Griffith, R., and Nesheim L. (2009), "A Micro-econometric Approach to Geographic Market Definition in Local Retail Markets: Demand Side Considerations", Uni-

versity College London, unpublished manuscript.

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

Blundell, R. and C. Meghir, (1987), "Bivariate alternatives to the Tobit model," Journal of Econometrics, 34, 179-200.

Chan, T. Y (2006), "Estimating a Continuous Hedonic-Choice Model With an Application to Demand for Soft Drinks," Rand Journal of Economics, 37(2), 466–482.

Crawford, G. and A. Yurukoglu (2012), The welfare effects of bundling in multichannel television markets," American Economic Review, 102, no. 2: 643-85.

Deaton, A. and J. Muellbauer (1980), Economics and Consumer Behavior, Cambridge: Cambridge University Press

Dubé, J.-P., (2004), "Multiple Discreteness and Product Differentiation: Demand for Carbonated Soft Drinks," Marketing Science, 23, 66-81.

Dubin, J. and D. McFadden, (1984), "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption," Econometrica, 52(2): 345-362.

Dubois, P., R. Griffith, and A. Nevo (2014) "Do Prices and Attributes Explain

International Differences in Food Purchases?" American Economic Review, 104(3): 832–867.

Elrod, T. and M. P. Keane (1995), "A Factor-Analytic Probit Model for Representing the Market Structure in Panel Data," Journal of Marketing Research, 32(1), 1-16.

Golan, A., J. M. Perloff, and E. Z. Shen (2001), "Estimating a Demand System with Nonnegativity Constraints: Mexican Meat Demand," Review of Economics and Statistics 83 541-550.

Gorman, W M. (1976), "Tricks with utility functions", in Artis, M. and A. R. Nobay (eds.), Essays in Economic Analysis, Cambridge University Press.

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

Heckman, J. (1979), "Sample Selection as a Specification Error," Econometrica, 47, 153–61.

Heien, D. and C. R. Wessells (1990), "Demand Systems Estimation with Microdata: A Censored Regression Approach," Journal of Business & Economic Statistics, 8(3), 365-371.

Hendel, I. (1999), "Estimating multiple-discrete choice models: An application to computerization returns," Review of Economic Studies, 66, 423–446.

Kim, J., G. M. Allenby, and P. E. Rossi (2007), "Product attributes and models of multiple discreteness," Journal of Econometrics, 138(1), 208-230.

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

Lee, L., and M. M. Pitt, (1986), "Microeconometric Demand Systems with Binding Non negativity Constraints: The Dual Approach," Econometrica, 54, 1237–42.

Levell, P., O'Connell, M. and Smith, K. (2017), "The exposure of households' food spending to tariff changes and exchange rate movements," IFS Briefing Note.

Lewbel, A. (1996), "Aggregation Without Separability: A Generalized Composite Commodity Theorem," American Economic Review, 86(3), 524-543.

Lewbel, A. and K. Pendakur (2009), "Tricks With Hicks: The EASI Demand System," American Economic Review, 99(3), 827-863.

Lewbel, A. and K. Pendakur (2017), "Unobserved Preference Heterogeneity in Demand Using Generalized Random Coefficients," Journal of Political Economy, 2017, 125(4) 1100-1148.

Meyerhoefer, C.D., C. K. Ranney, and D. E. Sahn (2005), "Consistent Estimation of Censored Demand Systems Using Panel Data," American Journal of Agricultural Economics, 87, 660-672.

Millimet, D. L. and R. Tchernis (2008), "Estimating high-Dimensional Demand Systems in the Presence of Many Binding Non-Negativity Constraints," Journal of Econometrics, 147(2), 384-395.

Nevo, A. (2000), "A Practitioner's Guide to Estimation of Random Coefficients Logit Models of Demand," Journal of Economics & Management Strategy, 9(4), 513-548.

Sam, A. G. and and Y. Zheng (2010), "Semiparametric Estimation of Consumer Demand Systems with Micro Data," American Journal of Agricultural Economics, 92, 246-257.

Shonkwiler, J. S., and S. T. Yen (1999), "Two-Step Estimation of a Censored System of Equations," American Journal of Agricultural Economics, 81, 972–82.

Thomassen, Ø. , Smith, H., Seiler, S., and P. Schiraldi (2017), Multi-category competition and market power: a model of supermarket pricing," American Economic Review, 107(8): 2308-51.

Tibshirani, R. (1996), "Regression shrinkage and selection via the lasso," Journal of the Royal Statistical Society B., 58(1), 267-288.

Van Soest, A., A. Kapteyn, and P. Kooreman (1993), "Coherency and regularity of demand systems with equality and inequality constraints," Journal of Econometrics, 57, 161-188.

Wales, T. J.,and A. D. Woodland (1983) "Estimation of Consumer Demand Systems with Binding Non-negativity Constraints," Journal of Econometrics, 21, 263–85.

Yen, S. T., and B. Lin, (2006) "A Sample Selection Approach to Censored Demand Systems," American Journal of Agricultural Economics, 88, 742–49.

Yen, S. T., B. Lin, and D. M. Smallwood (2003), "Quasi- and Simulated-likelihood Approaches to Censored Demand Systems: Food Consumption by Food Stamp Recipients in the United States," American Journal of Agricultural Economics, 85, 458–78.

A Tables

Table A.1: Most frequently purchased fruit categories

	Freq.	Pct.	Cum. Pct.
Banana	371,892	23.79	23.79
Apples	263,369	16.85	40.63
Grapes	156,189	9.99	50.63
Berries+Currants	152,731	9.77	60.40
Easy Peelers	135,073	8.64	69.04
Pears	91,062	5.82	74.86
Orange	62,599	4.00	78.86
Plums	50,879	3.25	82.12
Melons	41,845	2.68	84.80
Nectarines	37,954	2.43	87.22
Lemon	35,593	2.28	89.50
Kiwi Fruit	32,527	2.08	91.58
Pineapples	25,482	1.63	93.21
Avocado	20,810	1.33	94.54
Peaches	16,874	1.08	95.62
Grapefruit	15,248	0.98	96.60
Mango	15,096	0.97	97.56
Cherries	13,792	0.88	98.44
Lime	6,777	0.43	98.88
Dates	3,869	0.25	99.13
Apricot	3,349	0.21	99.34
Pomegranates	2,474	0.16	99.50
Sharon Fruit	2,059	0.13	99.63
Rhubarb	1,867	0.12	99.75
Passion Fruit	1,592	0.10	99.85
Paw-Paws	1,222	0.08	99.93
Lychees	1,114	0.07	100.00

Note: Using sample of all shopping trips in 2008, each row in the table records the frequency of purchase for a single category of fruit.

Table A.2: Number of categories purchased

	Freq.	Pct.	Cum. Pct.
1	377,096	48.18	48.18
2	200,632	25.63	73.81
3	108,527	13.86	87.67
4	53,301	6.81	94.48
5	24,929	3.18	97.67
6	10,889	1.39	99.06
7	4,590	0.59	99.64
8	1,756	0.22	99.87
9	643	0.08	99.95
10	234	0.03	99.98
11	96	0.01	99.99
12	45	0.01	100.00
13	11	0.00	100.00
14	10	0.00	100.00
15	2	0.00	100.00
Total	782,761	100.00	

Note: Using the sample of all shopping trips in 2008, the table records the frequency with which consumers purchased fruit baskets containing between 1 and 15 categories of fruit.

Table A.3: Most frequently purchased 2-item combinations

	Freq.	Pct.	Cum. Pct.
Banana, Apples	101533	25.03	25.03
Banana, Berries+Currants	52141	12.85	37.88
Banana, Easy Peelers	24442	6.03	43.91
Banana, Grapes	23977	5.91	49.82
Apples, Easy Peelers	18363	4.53	54.34
Berries+Currants, Apples	15931	3.93	58.27
Apples, Grapes	12052	2.97	61.24
Berries+Currants, Grapes	8592	2.12	63.36
Avocado, Banana	7915	1.95	65.31
Banana, Pears	7681	1.89	67.20
Apples, Pears	6299	1.55	68.76
Banana, Orange	5746	1.42	70.17
Berries+Currants, Easy Peelers	5506	1.36	71.53
Apples, Orange	5070	1.25	72.78
Easy Peelers, Grapes	4856	1.20	73.98
Banana, Melons	3551	0.88	74.85
Banana, Nectarines	3244	0.80	75.65
Banana, Lemon	3187	0.79	76.44
Banana, Kiwi Fruit	3144	0.78	77.21
Berries+Currants, Cherries	3018	0.74	77.96
Banana, Plums	2916	0.72	78.68
Avocado, Berries+Currants	2514	0.62	79.30
Banana, Cherries	2511	0.62	79.92
Berries+Currants, Melons	2151	0.53	80.45
Berries+Currants, Nectarines	2133	0.53	80.97
Apples, Kiwi Fruit	2043	0.50	81.48
Apples, Lemon	2009	0.50	81.97
Apples, Melons	1898	0.47	82.44
Banana, Grapefruit	1829	0.45	82.89
Apples, Nectarines	1803	0.44	83.33
Apples, Plums	1790	0.44	83.77
Avocado, Apples	1751	0.43	84.21
Grapes, Pears	1745	0.43	84.64
Easy Peelers, Pears	1734	0.43	85.06
Grapes, Orange	1508	0.37	85.44

Note: The table records the frequency with which various 2-item combinations were purchased.

Table A.4: Elasticities (1)

Price	Apricots	Avocados	Bananas	Berries	Cherries	Dates	Apples
$p_{Apricots}$	**-8.38**	-0.0049	0	0	0	0	0
$p_{Avocados}$	-0.191	**-2.33**	-0.00366	0.000848	-0.0232	0	0.0135
$p_{Bananas}$	0	-0.0253	**-0.237**	-0.212	-0.0895	-0.0204	-0.0364
$p_{Berries}$	0	0.00466	-0.168	**-1.45**	0.594	0.00516	-0.0105
$p_{Cherries}$	0	-0.00471	-0.00262	0.0219	**-4.61**	-0.000405	-0.00721
p_{Dates}	0	0	-0.00147	0.00047	-0.001	**-1.62**	-0.0535
p_{Apples}	0	0.0324	-0.0127	-0.00462	-0.0856	-0.257	**-0.395**
$p_{EasyPeelers}$	0	-0.012	-0.0749	0.0445	0.686	-0.0317	0.103
p_{Grapes}	0	0.0435	-0.121	0.1	0.0147	0.0625	-0.0135
$p_{Grapefruits}$	0	1.62e-06	-0.00277	6.76e-05	0	0.036	-0.00205
p_{Kiwis}	-3.28e-12	-0.0545	-0.0295	-0.106	0.0873	-0.00123	0.118
p_{Lemons}	0	0	0.00411	0.146	0.0293	0.00846	0.00845
p_{Limes}	0	0	-1.51e-05	0	0	0.000663	-1.05e-05
$p_{Lychees}$	0	0	0.00219	0	0	0.0332	0.000468
p_{Mangos}	0	0	-0.00195	-0.000596	0.0203	0.152	0.0248
p_{Melons}	0	0.259	-0.614	0.00996	0.229	-0.00969	0.0958
$p_{Nectarines}$	0.722	-0.0614	0.00882	-0.00412	0	-0.000843	0.0672
$p_{Oranges}$	0	0.00884	0.0991	0.0739	0.0722	0.148	-0.00488
$p_{Passionfruits}$	0	0	4.76e-05	0	0	0	0
$p_{Paw-paws}$	0	0	0.00316	0	0.000358	0.247	0.00391
$p_{Peaches}$	0	0	0.143	0.17	0.495	0.0678	-0.00136
p_{Pears}	0	0	0.000698	-0.000211	-0.0035	0.0703	0.0464
$p_{Pineapples}$	0	0.00672	-0.0159	0.119	0.345	0.000366	-0.00153
p_{Plums}	0	-0.00172	0.0121	-0.0191	0.0779	0.0572	-0.00483
$p_{Pomegranates}$	0	0	0.000482	7.42e-05	0	0.0418	0
$p_{Rhubarb}$	0	0	0	0	0	2.26e-05	0
$p_{Sharonfruits}$	0	0.00928	0	7.74e-06	0.00387	0	0

Note: Each column records the elasticities of demand of one fruit type with respect to its own price and the prices of the other fruits. The own-price elasticities are in boldface font.

Table A.5: Elasticities (2)

Price	Easy Peelers	Grapes	Grapefruits	Kiwis	Lemons	Limes	Lychees
$p_{Apricots}$	0	0	0	0	0	0	0
$p_{Avocados}$	-0.0022	0.0184	1.67e-05	-0.023	0	0	0
$p_{Bananas}$	-0.095	-0.354	-0.197	-0.086	0.0198	-0.0047	0.248
$p_{Berries}$	0.0448	0.233	0.00381	-0.244	0.561	0	-2.04e-11
$p_{Cherries}$	0.0255	0.00126	1.33e-11	0.00746	0.00415	0	0
p_{Dates}	-0.00291	0.0132	0.185	-0.00026	0.00296	0.0149	0.273
p_{Apples}	0.0453	-0.0137	-0.0506	0.12	0.0142	-0.00114	0.0185
$p_{EasyPeelers}$	**-0.317**	0.101	-0.122	-0.0258	0.0362	0.0241	-2.01e-12
p_{Grapes}	0.0437	**-1.74**	1.01	-0.000815	-0.0761	0.0329	0.888
$p_{Grapefruits}$	-0.00217	0.0415	**-4.06**	-0.000214	0.00141	0	0.152
p_{Kiwis}	-0.0112	-0.000818	-0.00522	**-1.07**	0.0248	0	0
p_{Lemons}	0.0095	-0.0461	0.0207	0.015	**-0.893**	0	0
p_{Limes}	9.8e-05	0.00031	0	0	0	**-2.35**	0
$p_{Lychees}$	0	0.0229	0.095	0	0	0	**-3.5**
p_{Mangos}	0.00295	0.00115	0.13	-0.00323	0.00227	0	0.0796
p_{Melons}	0.432	0.112	0.101	-0.0836	-1.67	0	0
$p_{Nectarines}$	-0.0124	-0.000783	0.00427	0.548	0.0338	0	0
$p_{Oranges}$	0.0103	1.51	1.01	-0.00023	-0.134	-0.0757	0.00427
$p_{Passionfruits}$	0	0	0.0071	0	0	0	0
$p_{Paw-paws}$	0.000625	-0.0045	0	0.000173	0	0	0
$p_{Peaches}$	-0.0469	-0.00914	0.375	-0.0101	-0.131	0.313	0.025
p_{Pears}	-0.00315	0.0169	-0.00504	-0.00645	0.0202	0.0221	0.00957
$p_{Pineapples}$	-0.017	0.0742	0.13	-0.00695	-0.00474	0	0
p_{Plums}	-0.000365	-0.0878	0.549	0.0179	0.139	-0.00332	0.0915
$p_{Pomegranates}$	0	0	-0.0167	0	0	0	0
$p_{Rhubarb}$	0	2.21e-05	0	0	0	0	0
$p_{Sharonfruits}$	0.000227	0	0	0.000514	0	0	0

Note: Each column records the elasticities of demand of one fruit type with respect to its own price and the prices of the other fruits. The own-price elasticities are in boldface font.

Table A.6: Elasticities (3)

Price	Mangos	Melons	Nectarines	Oranges	Passion fruits	Paw-paws	Peaches
$p_{Apricots}$	0	0	0.0346	0	0	0	0
$p_{Avocados}$	0	0.079	-0.115	0.00279	0	0	0
$p_{Bananas}$	-0.0492	-0.199	0.114	0.216	0.198	0.201	0.282
$p_{Berries}$	-0.0119	0.0174	-0.0423	0.128	0	0	0.266
$p_{Cherries}$	0.015	0.0141	0	0.00461	0	0.000668	0.0285
p_{Dates}	0.277	-0.00148	-0.000789	0.0233	0	1.14	0.00965
p_{Apples}	0.218	0.0703	0.302	-0.0037	0	0.0865	-0.000929
$p_{EasyPeelers}$	0.0586	0.154	-0.126	0.0176	0	0.0314	-0.0727
p_{Grapes}	0.00987	0.0805	-0.00346	1.12	4.06e-12	-0.0978	-0.00614
$p_{Grapefruits}$	0.0463	0.003	0.000778	0.031	0.415	0	0.0104
p_{Kiwis}	-0.028	-0.0605	2.43	-0.000172	0	0.00378	-0.00684
p_{Lemons}	0.0119	-0.00434	0.0907	-0.0607	0	0	-0.0534
p_{Limes}	0	0	0	-0.00053	0	0	0.00198
$p_{Lychees}$	0.0177	0	0	8.21e-05	0	0	0.000433
p_{Mangos}	**-1.04**	-0.00171	0.0251	-0.0173	0.216	0.136	0.000339
p_{Melons}	-0.0205	**-7.23**	-0.00801	0.165	0	0.0133	6.09
$p_{Nectarines}$	0.0489	-0.00131	**-4.18**	-0.00612	-0.0526	0.00313	-0.000685
$p_{Oranges}$	-0.201	0.16	-0.0363	**-2.42**	0.0484	1.08	0.281
$p_{Passionfruits}$	0.00131	0	-0.000164	2.54e-05	**-1.39**	0	0
$p_{Paw-paws}$	0.0538	0.00044	0.000636	0.037	0	**-4.88**	0.0067
$p_{Peaches}$	0.00435	0.652	-0.0045	0.311	0	0.217	**-1.8**
p_{Pears}	0.0947	-0.00116	0.0969	-0.00844	0	0.327	-0.00429
$p_{Pineapples}$	0.00116	0.06	0	0.00975	0	0	0.00993
p_{Plums}	0.0293	-0.000583	-0.0392	-0.0456	1.12e-11	0	0.0156
$p_{Pomegranates}$	0.000474	0	0	0	0	0	0
$p_{Rhubarb}$	0	0	0	0	0	0	0
$p_{Sharonfruits}$	0	0	0	0	0	0	0

Note: Each column records the elasticities of demand of one fruit type with respect to its own price and the prices of the other fruits. The own-price elasticities are in boldface font.

Table A.7: Elasticities (4)

Price	Pears	Pineapples	Plums	Pomegranates	Rhubarb	Sharon fruits
$p_{Apricots}$	0	0	0	0	0	0
$p_{Avocados}$	0	0.0124	-0.00154	0	0	5.6
$p_{Bananas}$	0.0112	-0.204	0.0748	5.88	0	0
$p_{Berries}$	-0.00269	1.2	-0.0941	0.718	0	0.0256
$p_{Cherries}$	-0.00165	0.129	0.0141	0	0	0.473
p_{Dates}	0.0817	0.000337	0.0256	36.9	0.26	0
p_{Apples}	0.259	-0.00675	-0.0104	0	0	0
$p_{EasyPeelers}$	-0.0399	-0.171	-0.00178	0	0	0.747
p_{Grapes}	0.0929	0.323	-0.186	-7.92e-11	1.2	0
$p_{Grapefruits}$	-0.00114	0.0232	0.0478	-2.86	0	0
p_{Kiwis}	-0.0355	-0.034	0.0379	0	0	0.735
p_{Lemons}	0.0669	-0.0125	0.179	0	0	0
p_{Limes}	0.00114	0	-6.6e-05	0	0	0
$p_{Lychees}$	0.00135	0	0.00499	0	0	0
p_{Mangos}	0.0603	0.000584	0.00719	0.229	0	0
p_{Melons}	5.9	0.36	-0.00158	0	0	0
$p_{Nectarines}$	0.12	0	-0.0188	0	0	0
$p_{Oranges}$	-0.0621	0.0568	-0.129	0	0	0
$p_{Passionfruits}$	0	0	0	0	0	0
$p_{Paw-paws}$	0.0826	0	0	0	0	0
$p_{Peaches}$	-0.035	0.0642	0.049	0	0	0
p_{Pears}	**-0.881**	-0.191	0.127	0	0	0
$p_{Pineapples}$	-0.241	**-3.02**	0.301	0	0	0
p_{Plums}	0.33	0.621	**-0.936**	-1.4	0	0.309
$p_{Pomegranates}$	0	0	-0.000708	**-41.1**	0	0
$p_{Rhubarb}$	0	0	0	0	**-2.63**	0
$p_{Sharonfruits}$	0	0	0.000458	0	0	**-10.6**

Note: Each column records the elasticities of demand of one fruit type with respect to its own price and the prices of the other fruits. The own-price elasticities are in boldface font.

Table A.8: Percentage change in price due to tax/price change

Fruit	Baseline	Scenario 1 EU tariff	Scenario 2 UK cost shock	Scenario 3 Merger	Scenario 4 Subsidy	Scenario 5 VAT
Apricots	2.44	6.52%	3.33e-13%	2.21%	-10%	20%
Avocados	4.4	1.93%	2.54e-11%	1.73%	-10%	20%
Bananas	1.05	1.69e-11%	1.69e-11%	1.56%	-10%	20%
Berries	6	3.54%	4.3%	1.75%	-10%	20%
Cherries	7.1	3.47%	1.9%	1.45%	-10%	20%
Dates	1.54	0.05%	-1.95e-11%	1.08%	-10%	20%
Apples	1.41	2.86%	1.54%	1.66%	-10%	20%
Easy Peelers	1.78	4.24%	-2.15e-11%	1.71%	-10%	20%
Grapes	2.34	3.39%	0.08%	1.46%	-10%	20%
Grapefruits	0.886	0.87%	1.11e-11%	1.75%	-10%	20%
Kiwis	1.34	6.65%	7.97e-12%	1.55%	-10%	20%
Lemons	1.91	5.21%	-5.44e-12%	1.7%	-10%	20%
Limes	0.892	2.27%	-1.64e-11%	1.91%	-10%	20%
Lychees	5.37	5.17e-12%	5.17e-12%	1.4%	-10%	20%
Mangos	1.37	7.73e-12%	7.73e-12%	1.38%	-10%	20%
Melons	1.1	1.81%	3.49e-11%	1.51%	-10%	20%
Nectarines	2.17	5.96%	3.64e-11%	1.91%	-10%	20%
Oranges	1.44	2.88%	1.23e-11%	1.66%	-10%	20%
Passion fruits	1.8	-1.99e-11%	-1.99e-11%	1.51%	-10%	20%
Paw-paws	3.23	2.12e-11%	2.12e-11%	1.48%	-10%	20%
Peaches	2.06	8.97%	-1.98e-12%	1.54%	-10%	20%
Pears	1.49	5.15%	1.54%	1.59%	-10%	20%
Pineapples	0.985	2.01e-11%	2.01e-11%	1.22%	-10%	20%
Plums	2.15	3.19%	0.26%	1.61%	-10%	20%
Pomegranates	1.77	3.71%	-1.32e-11%	1.14%	-10%	20%
Rhubarb	4.51	0.11%	9.87%	1.78%	-10%	20%
Sharon fruits	9.75	0.5%	2.19e-11%	1.7%	-10%	20%

Note: The first column shows the baseline price for each fruit (GBP per kilogram). The remaining columns show the percentage impact of the change in tax or prices.

Table A.9: Percentage change in demand due to tax/price change

Fruit	Baseline (kg)	Scenario 1 EU tariff	Scenario 2 UK cost shock	Scenario 3 Merger	Scenario 4 Subsidy	Scenario 5 VAT
Apricots	0.00177	-29%	1.87e-11%	-13%	116%	-72%
Avocados	0.0384	-4.39%	0.0592%	-3.69%	26.4%	-31.8%
Bananas	1.11	0.657%	-0.676%	-0.741%	5.08%	-9.22%
Berries	0.155	-2.7%	-7.61%	-2.12%	12%	-19.3%
Cherries	0.00482	-7.57%	-5.32%	-2.92%	28.7%	-45%
Dates	0.0548	1.18%	-0.232%	-0.836%	10.8%	-16.4%
Apples	0.288	0.972%	-0.933%	-0.0613%	0.291%	-1.08%
Easy Peelers	0.518	-1.25%	0.468%	-0.405%	2.39%	-4.6%
Grapes	0.17	-0.428%	0.72%	0.258%	0.973%	-3.23%
Grapefruits	0.0186	12.5%	0.707%	-2.22%	8.73%	-12.3%
Kiwis	0.299	-4.59%	-0.737%	-1.16%	9.13%	-14.9%
Lemons	0.127	-3.77%	4.01%	-0.402%	3.67%	-7.68%
Limes	0.00421	-2.93%	0.0341%	-4.01%	34.7%	-37.8%
Lychees	0.00192	4.44%	0.138%	-2.17%	20.5%	-34%
Mangos	0.0338	1.11%	0.531%	-0.645%	5.22%	-9.97%
Melons	0.505	6.25%	0.122%	-1.12%	7.93%	-14.3%
Nectarines	0.0416	-8.88%	3.01%	-3.62%	16.5%	-26.3%
Oranges	0.372	0.0718%	0.918%	-1.22%	5.48%	-8.57%
Passion fruits	0.000156	0.187%	3.6e-11%	-0.782%	5.63%	-9.77%
Paw-paws	0.00569	4.64%	0.66%	-2.68%	19.4%	-26.1%
Peaches	0.288	-18.3%	1.2%	-0.587%	4.35%	-8.33%
Pears	0.0489	-3.08%	-1.12%	-0.338%	2%	-3.43%
Pineapples	0.0933	19.9%	15.6%	0.533%	9.31%	-14%
Plums	0.0879	-1.96%	-1.87%	-1.1%	5.53%	-10.3%
Pomegranates	5.41e-05	-35.6%	2.73%	-3.43%	15.6%	-31.2%
Rhubarb	1.63e-06	3.79%	-100%	-2.65%	-100%	-23.4%
Sharon fruits	2.87e-05	15.7%	1.09%	-7.37%	12.9%	-34.7%

Note: The first column shows baseline demand for each fruit (kilograms per household per shopping trip). The remaining columns show the percentage change in demand resulting from the change in tax or prices.

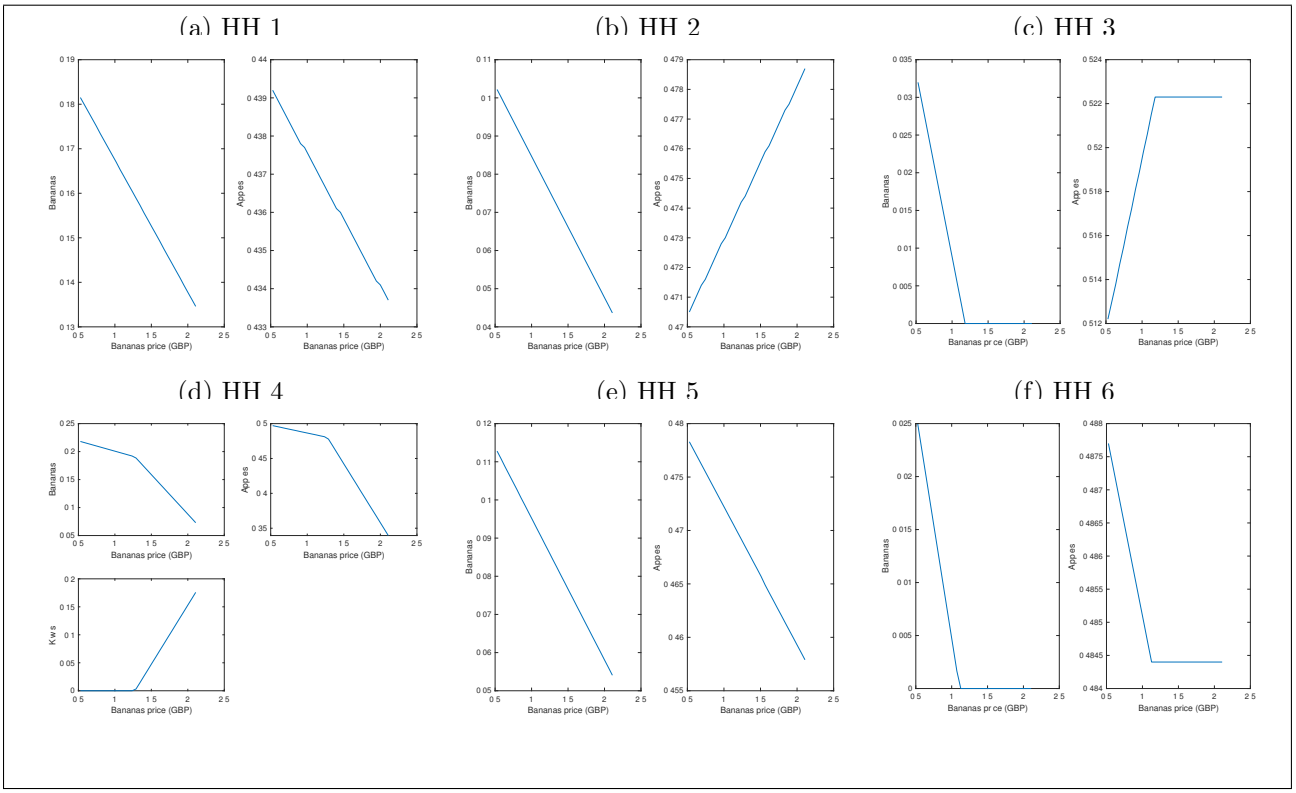
Table A.10: Tax impact on expenditure and welfare

	Baseline	Scenario 1 EU tariff	Scenario 2 UK cost shock	Scenario 3 Merger	Scenario 4 Subsidy	Scenario 5 VAT
Consumer expenditure						
10th percentile	1.1	0.85%	-0.0579%	0.319%	-4.04%	5.2%
25th percentile	2.15	0.914%	0.306%	0.555%	-3.46%	6.4%
50th percentile	4.41	0.982%	0.0301%	0.509%	-3.59%	5.88%
75th percentile	9.39	0.765%	-0.17%	0.413%	-4.03%	5.73%
90th percentile	16.4	1.95%	-0.147%	0.842%	-4.75%	8.23%
Change in consumer surplus (GBP)						
10th percentile	2.68	-0.0487	-0.0127	-0.0209	0.196	-0.325
25th percentile	7.14	-0.0905	-0.0162	-0.0559	0.343	-0.598
50th percentile	16.3	-0.151	-0.0163	-0.0649	0.598	-1.02
75th percentile	31.7	-0.259	-0.0331	-0.134	0.917	-1.75
90th percentile	53.3	-0.437	-0.076	-0.204	1.46	-2.61
Per capita effects						
Consumer surplus (GBP)	23.4	-0.228	-0.0472	-0.112	0.721	-1.32
Tax revenue (GBP)	0.0	0.222	0.0456	0.111	-0.744	1.24
Firm Revenue	6.98	-0.13	-0.0444	-0.0683	0.457	-0.761

Note: The first column shows the baseline values for expenditure, consumer surplus, firm revenue and tax revenue. All amounts are measured in pounds per household per shopping trip. Columns 2 - 7 show the percentage change in expenditure, the absolute change in consumer surplus, the absolute change in firm revenue and the absolute change in tax revenue arising in each scenario. Because of quasilinear utility the change in consumer surplus equals compensating variation.

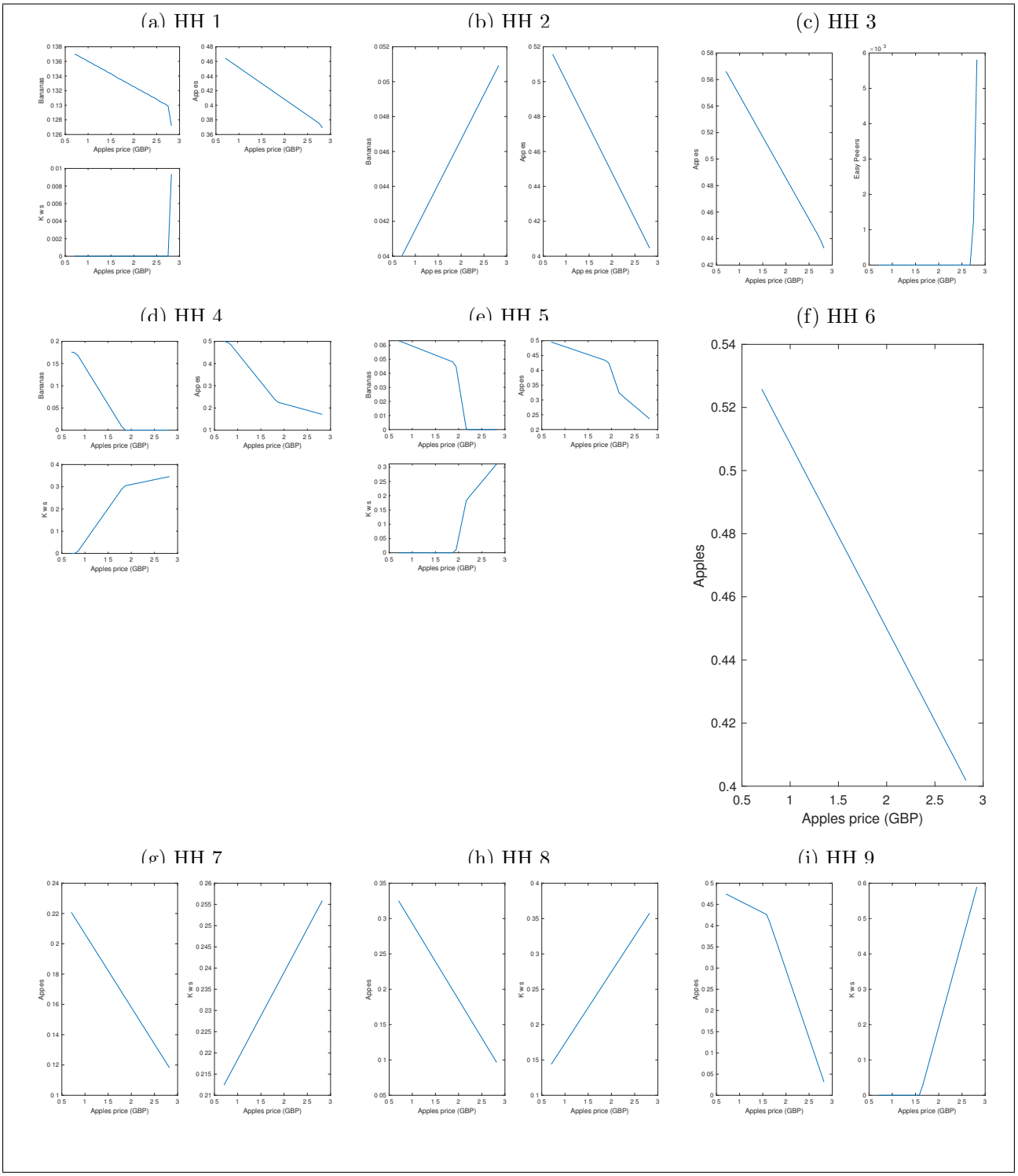
B Figures

Figure B.1: Demand vs. banana price: by household type



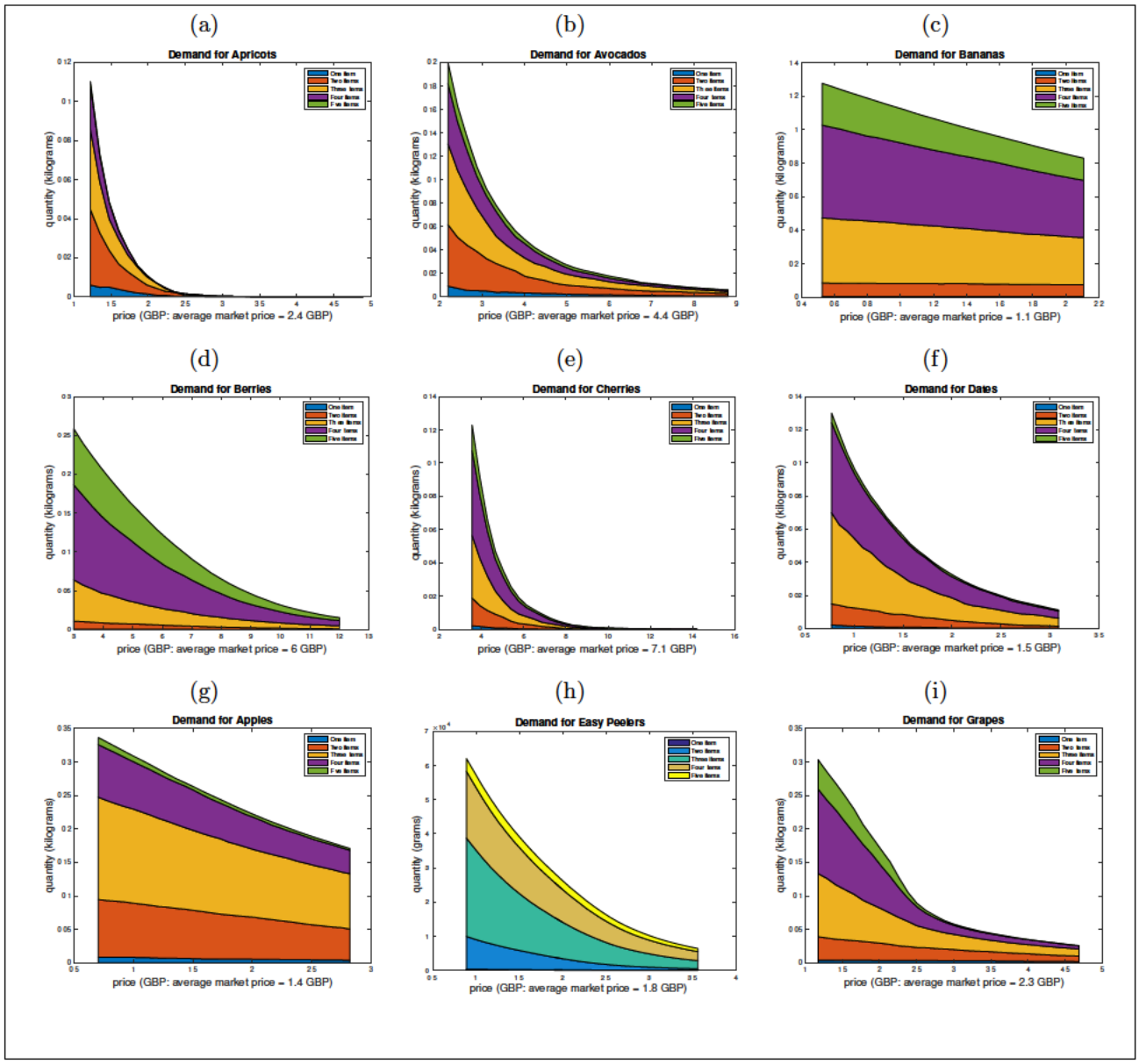
Each panel shows demand for fruit vs. banana price for one household type. Only fruits with non-zero demand are shown.

Figure B.2: Demand vs. apple price: by household type



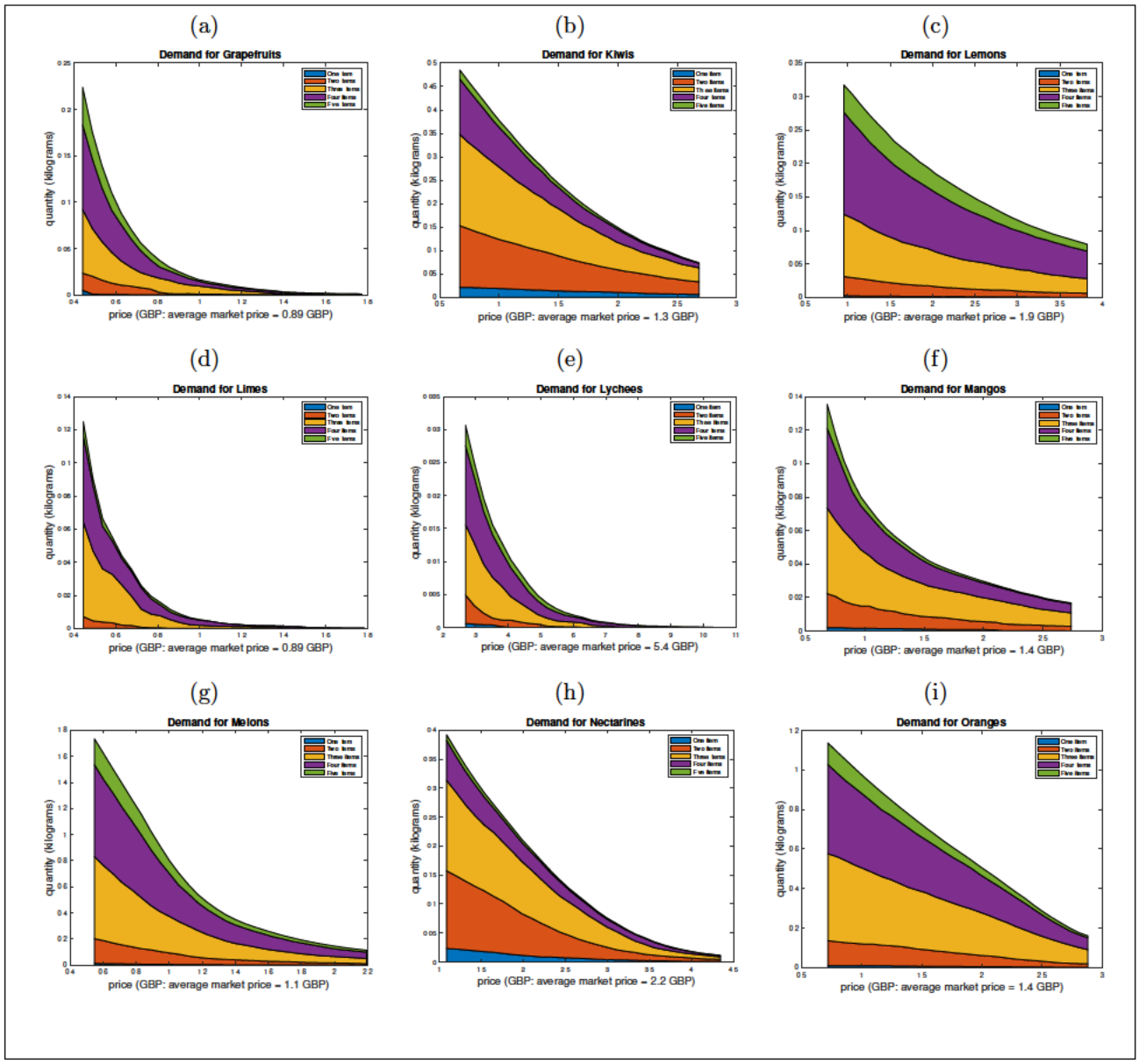
Each panel shows demand for fruit vs. apple price for one household type. Only fruits with non-zero demand are shown.

Figure B.3: Aggregate demand curves (1)



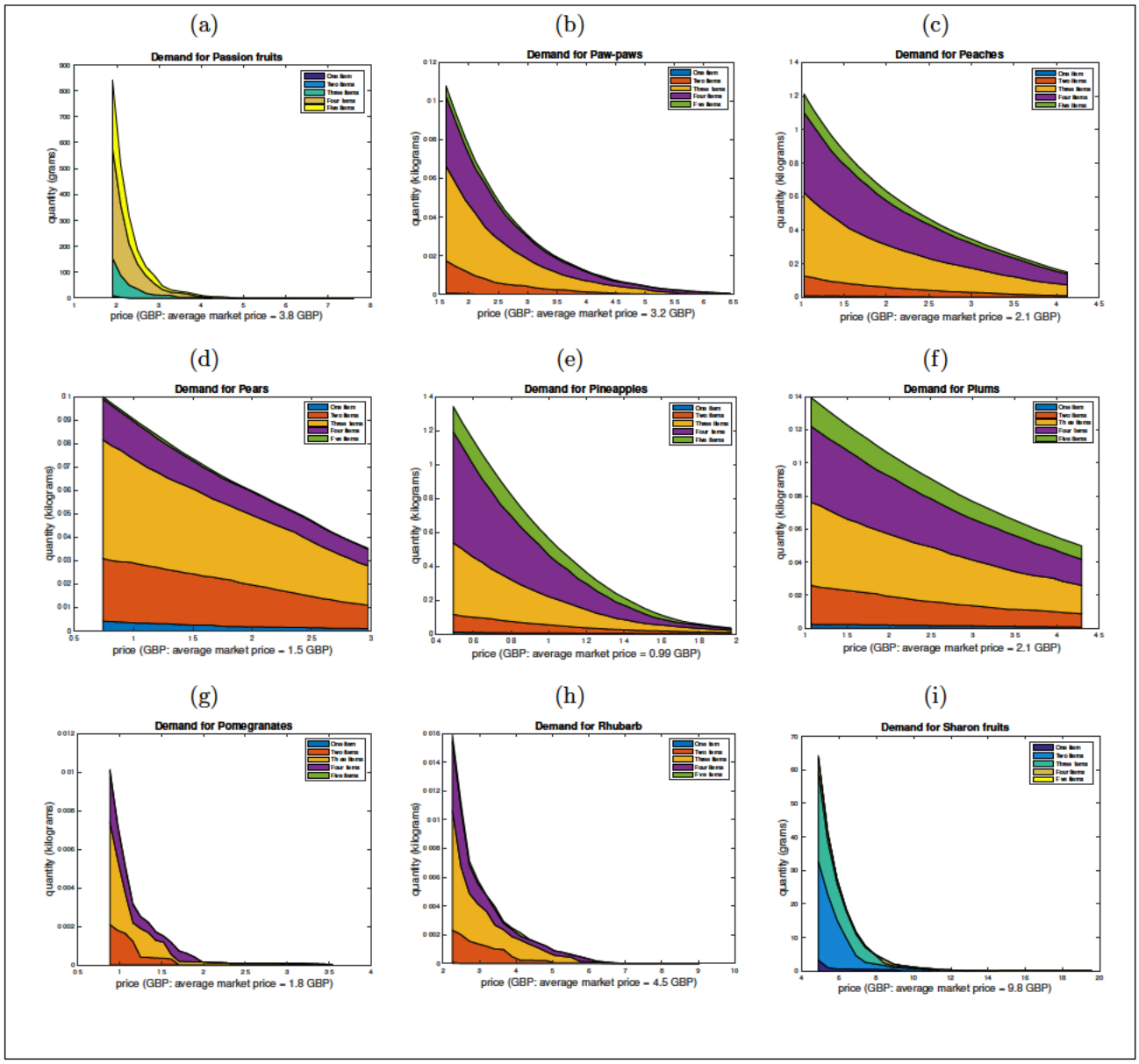
Each panel shows aggregate demand for a fruit category as that fruit's price varies from 50% to 200% of its baseline level. The figure also shows what fraction of demand comes from baskets of size 1 through 5.

Figure B.4: Aggregate demand curves (2)



Each panel shows aggregate demand for a fruit category as that fruit's price varies from 50% to 200% of its baseline level. The figure also shows what fraction of demand comes from baskets of size 1 through 5.

Figure B.5: Aggregate demand curves (3)



Each panel shows aggregate demand for a fruit category as that fruit's price varies from 50% to 200% of its baseline level. The figure also shows what fraction of demand comes from baskets of size 1 through 5.

Supplementary Appendix to "Sparse demand systems: corners and complements"

Arthur Lewbel

Boston College and IFS

and

Lars Nesheim

CeMMAP, IFS and UCL

December 2019

Abstract

This Supplementary Appendix presents technical details for the paper "Sparse demand systems: corners and complements." These include details of the hyperspherical transformation, the log likelihood function, and the hedonic price estimation. It also presents some summary statistics for the data.

Correspondence: l.nesheim@ucl.ac.uk

Acknowledgement: We gratefully acknowledge financial support from the ESRC through the ESRC Centre for Microdata Methods and Practice (CeMMAP) grant number ES/I034021/1 and the European Research Council (ERC) under ERC-2009-AdG grant agreement number 249529. Data supplied by Kantar Worldpanel. The use of Kantar Worldpanel data in this work does not imply the endorsement of Kantar Worldpanel in relation to the interpretation or analysis of the data. All errors and omissions remained the responsibility of the authors.

1 Introduction

This Supplementary Appendix presents technical details for the paper "Sparse demand systems: corners and complements." Appendix A describes the hyperspherical coordinate representation used in the paper. Appendix B derives the log likelihood function and presents algebraic manipulations that are used to compute the value of the log likelihood. Appendix C presents additional summary statistics for the data. Appendix D presents details of the hedonic price functions estimated.

A Hyperspherical representation of B

As discussed in Section 5.2 in the paper, it is convenient to reparameterize the matrix B in hyperspherical coordinates. This representation is derived as follows. Since B is upper triangular, $b_{kj} = 0$ if $k > j$. The number of nonzero elements in column B_j is $\overline{k} = \min\{K, j\}$. Let $C_j = \left[c_{1j}, ..., c_{\overline{k}-1}\right]^T$. The hyperspherical coordinate representation of the nonzero elements of B_j is given by $(d_j, C_j) = H(B_j)$ where H^{-1} is defined by

$$\begin{aligned}
B(1,j) &= d_j \cos(c_{1j}) \\
B(2,j) &= d_j \sin(c_{1j}) \cos(c_{2j}) \\
B(3,j) &= d_j \sin(c_{1j}) \sin(c_{2j}) \cos(c_{3j}) \\
&\vdots \\
B\left(\overline{k}-1,j\right) &= d_j \sin c_{1j} \cdots \sin\left(c_{\overline{k}-2}\right) \cos\left(c_{\overline{k}-1}\right) \\
B\left(\overline{k},j\right) &= d_j \sin(c_{1j}) \cdots \sin\left(c_{\overline{k}-2}\right) \sin\left(c_{\overline{k}-1}\right)
\end{aligned} \tag{A.1}$$

with $d_j > 0$, $c_{kj} \in [0, \pi]$ for $k < \overline{k} - 2$ and $c_{\overline{k}-1} \in [0, 2\pi)$.

B Estimation details

In this section we derive the components of the log likelihood function for 3 cases. Case 1 applies to observations in which a household purchased K goods. Case 2 applies to observations in which a household bought more than zero and fewer than K goods. Case 3 applies to observations in which a household bought zero goods.

B.1 Case 1: choice of K goods

The notation is the same as the main paper as defined in Section 3 and in Section 5.2

We drop household subscripts h to ease notation.

Suppose the goods are sorted so that $q = (q_1, 0)$. Let $p = (p_1, p_2)$ be the corresponding vector of prices. That is, the first K elements are non-negative and the remaining $J - K$ elements are 0. Let $B = [B_1\ B_2]$ as in Section 3.2.

Inverting the demand function given in equation (3.6) in Section 3.2 in the paper, inverse demand is

$$\begin{aligned} e &= \left(B_1^T\right)^{-1}\left(p_1 + B_1^T B_1 q_1\right) \\ &= \left(B_1^T\right)^{-1} p_1 + B_1 q_1 \\ p_2 &\geq B_2^T\left(B_1^T\right)^{-1} p_1. \end{aligned}$$

Since B is a function of η, $\eta \sim N(0, I)$ and $e \sim N(\mu, \Sigma)$, the case 1 log-likelihood is

$$\ln f_1(q, p, \theta) = \int_\eta \left\{\ln \phi\left[\left(B_1^T\right)^{-1} p_1 + B_1 q_1, \mu, \Sigma\right] + \ln\left(\det\left(B_1\right)\right)\right\} \phi(\eta, 0, I) d\eta$$

where f_1 is the case 1 density of q conditional on p and ϕ is the normal density function. Note that parameter values must satisfy the constraints that $p_2 \geq B_2^T\left(B_1^T\right)^{-1} p_1$.

B.2 Case 2: Choice of fewer than K goods

We first derive the likelihood function for fixed B.

Suppose a household chooses $q = (q_1, 0)$ with $q_1 > 0$ and $\dim(q_1) = d_1 < K$. In this case, for each q_1, there are multiple vectors e that satisfy the first order conditions

$$-p_1 - B_1^T (B_1 q_1 - e) = 0 \tag{B.1}$$

$$-p_2 - B_2^T (B_1 q_1 - e) \leq 0 \tag{B.2}$$

$$q_1 > 0. \tag{B.3}$$

In fact, the set of e values satisfying the first order conditions is a linear space of dimension $K - d_1$. In these expressions, B_1 is a $K \times d_1$ matrix with $d_1 < K$ and B_2 is a $(K \times J - d_1)$ matrix.

Let

$$B_1 = USV^T$$

be the singular value decomposition of B_1 where U is orthogonal $(K \times K)$, $S = \begin{bmatrix} S_1 \\ 0 \end{bmatrix}^T$ where S_1 is diagonal $(d_1 \times d_1)$ and V is orthogonal $(d_1 \times d_1)$. Define $\widetilde{e} = U^T e$ and partition $\widetilde{e} = (\widetilde{e}_1, \widetilde{e}_2)$ where $\widetilde{e}_1$ is $(d_1 \times 1)$ and $\widetilde{e}_2$ is $(d_2 \times 1)$. Then rewrite $(B.1)$ as

$$V \begin{bmatrix} S_1 & 0 \end{bmatrix} \begin{bmatrix} \widetilde{e}_1 \\ \widetilde{e}_2 \end{bmatrix} = p_1 + B_1^T B_1 q_1$$

or

$$V S_1 \widetilde{e}_1 = p_1 + B_1^T B_1 q_1. \tag{B.4}$$

For each q_1 there are multiple vectors $\widetilde{e}$ that solve $(B.4)$. In fact, there is a linear space

of dimension d_2. In other words, for each $(q_1, \widetilde{e}_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, there is a unique $\widetilde{e}_1$ defined by

$$\widetilde{e}_1 = G_0 p_1 + G_1 q_1 \tag{B.5}$$

where

$$\begin{aligned} G_0 &= S_1^{-1} V^T \\ G_1 &= S_1^{-1} V^T \left(B_1^T B_1 \right). \end{aligned} \tag{B.6}$$

Since B_1 has rank d_1 by assumption, S_1 is a $(d_1 \times d_1)$ invertible diagonal matrix and by construction $V^{-1} = V^T$.

Since

$$\widetilde{e} = U^T e$$

$\widetilde{e} \sim N\left(\widetilde{\mu}, \widetilde{\Sigma}\right)$ where $\widetilde{\mu} = U^T \mu$ and $\widetilde{\Sigma} = U^T \Sigma U$. Consider the partially observed random vector $(q_1, \widetilde{e}_2)$. q_1 is observed but $\widetilde{e}_2$ is not. The expressions above imply that the density of $(q_1, \widetilde{e}_2)$ is

$$f_{q_1 \widetilde{e}_2}(q_1, \widetilde{e}_2) = f_{\widetilde{e}}(G_0 p_1 + G_1 q_1, \widetilde{e}_2) \cdot \det(G_1)$$

where (G_0, G_1) are defined in $(B.6)$.

We observe q_1 if inequality $(B.2)$ is satisfied. Since $B_1 = USV^T$ and $e = U\widetilde{e}$, this is equivalent to

$$-p_2 - B_2^T U \left(SV^T q_1 - \widetilde{e} \right) \leq 0. \tag{B.7}$$

Partitioning $\widetilde{B}_2 = U^T B_2$ $(K \times J - d_1)$ as

$$\widetilde{B}_2 = \begin{bmatrix} \widetilde{B}_{21} \\ \widetilde{B}_{22} \end{bmatrix}$$

where $\widetilde{B}_{21}$ is size $(d_1 \times J - d_1)$ and $\widetilde{B}_{22}$ is size $(d_2 \times J - d_1)$, inequality $(B.7)$ is

$$-p_2 - \left[\begin{array}{cc} \widetilde{B}_{21}^T & \widetilde{B}_{22}^T \end{array} \right] \left(\left[\begin{array}{c} S_1 V^T q_1 \\ 0 \end{array} \right] - \left[\begin{array}{c} \widetilde{e}_1 \\ \widetilde{e}_2 \end{array} \right] \right) \leq 0$$

or

$$-p_2 - \widetilde{B}_{21}^T \left(S_1 V^T q_1 - \widetilde{e}_1 \right) + \widetilde{B}_{22}^T \widetilde{e}_2 \leq 0$$

Substituting from equation $(B.5)$ this is equivalent to

$$\widetilde{B}_{22}^T \widetilde{e}_2 \leq p_2 - \widetilde{B}_{21}^T G_0 p_1 + \widetilde{B}_{21}^T \left(S_1 V^T - G_1 \right) q_1. \tag{B.8}$$

Rewrite $(B.8)$ as

$$M_1 \widetilde{e}_2 \leq M_2$$

where

$$M_1 = \widetilde{B}_{22}^T$$

is a $(J - d_1 \times d_2)$ matrix and

$$M_2 = p_2 - \widetilde{B}_{21}^T G_0 p_1 + \widetilde{B}_{21}^T \left(S_1 V^T - G_1 \right) q_1$$

is $(J - d_1 \times 1)$.

Then the Case 2 likelihood, conditional on $B(\eta)$ and p is

$$f_2 \left[q, p, B(\eta), \theta \right] = \int f_{q_1 \widetilde{e}_2} \left(q_1, \widetilde{e}_2 \right) 1 \left(M_1 \widetilde{e}_2 \leq M_2 \right) d\widetilde{e}_2. \tag{B.9}$$

Note that $f_2 \left[q, p, B(\eta), \theta \right] = 0$ if $\Pr \left(M_1 \widetilde{e}_2 \leq M_2 \right) = 0$.

Let $d_2 = K - d_1$, let $\widetilde{\Sigma}_{22} = \widetilde{C}_2^T C_2$ be the variance of $\widetilde{e}_2$. That is $\widetilde{C}_2^T$ is the upper triangular cholesky decomposition of $\widetilde{\Sigma}_{22}$. Define $\widetilde{e}_2 = \widetilde{C}_2^T z_2 + \widetilde{\mu}_2$ and note that after a change of variables

the density of $\widetilde{e}$ can be written

$$f_{\widetilde{e}}\left(\widetilde{e}_1, z_2\right) = f_{\widetilde{e}_1}\left(\widetilde{e}_1, \nu_1\left(z_2\right), \Omega_1\right) \frac{e^{-0.5 z_2^T z_2}}{(2\pi)^{\frac{d_2}{2}}}$$

where $\widetilde{e}_1 \sim N\left(\nu_1, \Omega_1\right)$ and $z_2 \sim N\left(0, I\right)$ where

$$\begin{aligned} v_1 &= \widetilde{\mu}_1 + \widetilde{\Sigma}_{12}\widetilde{C}_2^{-1} z_2 \\ \Omega_1 &= \widetilde{\Sigma}_{11} - \widetilde{\Sigma}_{12}\widetilde{\Sigma}_{22}^{-1}\widetilde{\Sigma}_{21}. \end{aligned}$$

Therefore, $(B.9)$ can be written

$$f_2\left[q, p, B(\eta), \theta\right] = \int f_{q_1 z_2}\left(q_1, z_2\right) 1\left(\widetilde{M}_1 z_2 \leq \widetilde{M}_2\right) dz_2 \tag{B.10}$$

where

$$\begin{aligned} f_{q_1 z_2}\left(q_1, z_2\right) &= f_{e_1|z_2}\left(G_0 + G_1 q_1, \nu_1\left(z_2\right), \Omega_1\right) \frac{e^{-0.5 z_2^T z_2}}{(2\pi)^{\frac{d_2}{2}}} \\ &= \widetilde{f}_{q_1 z_2}\left(q_1, z_2\right) \frac{e^{-0.5 z_2^T z_2}}{(2\pi)^{\frac{d_2}{2}}} \end{aligned}$$

$$\begin{aligned} \widetilde{M}_1 &= M_1 \widetilde{C}_2^T \\ \widetilde{M}_2 &= M_2 - M_1 \widetilde{\mu}_2 \end{aligned}$$

The matrix $\widetilde{M}_1$ has the QR decomposition

$$\widetilde{M}_1 = RQ$$

where R is $(J - d_1 \times d_2)$ lower triangular and Q is $(d_2 \times d_2)$ orthogonal. Then using the

change of variable $z_2 = Q^{-1}x$, the integral can be written as

$$f_2\left[q,p,B(\eta),\theta\right] = \int\limits_{RQz_2 \le D} \widetilde{f}_{q_1 z_2}\left(q_1,z_2\right)\frac{e^{-0.5z_2^T z_2}}{(2\pi)^{\frac{d_2}{2}}}dz_2 \tag{B.11}$$

$$= \int\limits_{Rx \le D} \widetilde{f}_{q_1 z_2}\left(q_1,Q^{-1}x\right)\frac{e^{-0.5x^T x}}{(2\pi)^{\frac{d_2}{2}}}dx \tag{B.12}$$

since Q is an orthogonal matrix. (That is $Q^{-1}Q = I$ and $\det(Q) = 1$) The matrix R is lower triangular. Therefore, row i has at most i nonzero elements.

Start from x_{d_2}. Let $J_{d_2}^+$ be the set of rows of R that have positive elements in column d_2 and $J_{d_2}^-$ the set with negative elements. Then for all $j \in J_{d_2}^+$,

$$-\infty \le x_{d_2} \le \frac{D_j - \sum\limits_{i<d_2} R\left(j,i\right)x_i}{R\left(j,d_2\right)}$$

and for all $j \in J_{d_2}^-$,

$$\frac{D_j - \sum\limits_{i<d_2} R\left(j,i\right)x_i}{R\left(j,d_2\right)} \le x_{d_2} \le \infty.$$

So, the bounds on x_{d_2} are $x_{d_2} \in \left[x_{d_2}^L, x_{d_2}^H\right]$ where

$$x_{d_2}^L = \max\left(-\infty, \max_{j \in J_{d_2}^-}\left(\frac{D_j - \sum\limits_{i<d_2} R\left(j,i\right)x_i}{R\left(j,d_2\right)}\right)\right)$$

and

$$x_{d_2}^H = \min\left(\infty, \min_{j \in J_{d_2}^+}\left(\frac{D_j - \sum\limits_{i<d_2} R\left(j,i\right)x_i}{R\left(j,d_2\right)}\right)\right).$$

We repeat the calculation for $j = d_2 - 1$ through 1. Then the integral is

$$f_2\left[q, p, B(\eta), \theta\right] = \int_{x_1^L}^{x_1^H} \cdots \int_{x_{d_2}^L}^{x_{d_2}^H} \widetilde{f}_{q_1 z_2}\left(q_1, Q^{-1}x\right) \frac{e^{-0.5x^T x}}{(2\pi)^{\frac{d_2}{2}}} dx. \tag{B.13}$$

Next for all $j \leq d_2$ define $u_j = \Phi\left(x_j\right)$. Then making the change of variables, the integral is equivalent to

$$f_2\left[q, p, B(\eta), \theta\right] = \int_{u_1^L}^{u_1^H} \cdots \int_{u_{d_2}^L}^{u_{d_2}^H} \widetilde{f}_{q_1 z_2}\left(q_1, Q^{-1}x\left(u\right)\right) du \tag{B.14}$$

where

$$\begin{aligned} u_j^L &= \Phi\left(x_j^L\right) \\ u_j^H &= \Phi\left(x_j^H\right). \end{aligned}$$

Finally, for all $j \leq d_2$ making the change of variable $u_j = \frac{\left(u_j^H - u_j^L\right)(1+v_j)}{2}$, this is equivalent to

$$f_2\left[q, p, B(\eta), \theta\right] = \int_{-1}^{1} \cdots \int_{-1}^{1} \prod_{j=1}^{d_2} \left(\frac{u_j^H - u_j^L}{2}\right) \widetilde{f}_{q_1 z_2}\left(q_1, Q^{-1}x\left(v\right)\right) dv. \tag{B.15}$$

This equals 0 if $u_j^H \leq u_j^L$ for any j.

The conditional density function f_2 depends on the parameters θ and on the random coefficient η. Integrating out the random coefficients, the Case 2 likelihood function is

$$\ln f_2\left(q, p, \theta\right) = \int_{\eta} f_2\left[q, p, B(\eta), \theta\right] \phi(\eta) d\eta.$$

B.3 Case 3: Choice of 0 goods

Suppose a household chooses $q = 0$. In this case, the first-order conditions are

$$-p + B^T e \leq 0. \tag{B.16}$$

In this inequality, B is a $K \times J$ matrix. Rewrite the inequality as

$$B^T e \leq p. \tag{B.17}$$

Let $e = Cz + \mu$. Then this is equivalent to

$$B^T (Cz + \mu) \leq p$$

$$\widetilde{B}^T z \leq p - B^T \mu.$$

where $\widetilde{B} = C^T B$. Let

$$\widetilde{B} = QR$$

be the QR decomposition of $\widetilde{B}$ where R is $(K \times J)$ lower triangular. Since Q is orthogonal $Q^T Q = I$ and $\det(Q) = 1$.

Then defining $z = Qx$, the likelihood conditional on $B(\eta)$ and p can be written

$$f_3[q, p, B(\eta), \theta] = \int_{R^T x \leq p - B^T \mu} \frac{e^{-0.5 x^T x}}{(2\pi)^{\frac{K}{2}}} dx. \tag{B.18}$$

Start from x_K. Let J_K^+ be the set of rows of C that have positive elements in column K and J_K^- the set with negative elements. Let $D = p - B^T \mu$. Then for all $j \in J_K^+$,

$$-\infty \leq x_K \leq \frac{D_j - \sum_{i<K} R(j,i)\, x_i}{R(j,K)}$$

and for all $j \in J_K^-$,

$$\frac{D_j - \sum_{i<K} R(j,i)\, x_i}{R(j,K)} \leq x_K \leq \infty.$$

So, the bounds on x_K are $x_K \in \left[x_K^L, x_K^H\right]$ where

$$x_K^L = \max\left(-\infty, \max_{j \in J_{d_2}^-}\left(\frac{D_j - \sum_{i<K} R(j,i)\, x_i}{R(j,K)}\right)\right)$$

and

$$x_K^H = \min\left(\infty, \min_{j \in J_K^+}\left(\frac{D_j - \sum_{i<K} R(j,i)\, x_i}{R(j,K)}\right)\right).$$

We repeat the calculation for $j = K-1$ through 1. Then the integral is

$$f_3\left[q, p, B(\eta), \theta\right] = \int_{x_1^L}^{x_1^H} \cdots \int_{x_K^L}^{x_K^H} \frac{e^{-0.5x^T x}}{(2\pi)^{\frac{d_2}{2}}} dx. \tag{B.19}$$

Next for all $j \leq K$ define $u_j = \Phi\left(x_j\right)$. Then making the change of variables, the integral is equivalent to

$$f_3\left[q, p, B(\eta), \theta\right] = \int_{u_1^L}^{u_1^H} \cdots \int_{u_K^L}^{u_K^H} du \tag{B.20}$$

where

$$\begin{aligned} u_j^L &= \Phi\left(x_j^L\right) \\ u_j^H &= \Phi\left(x_j^H\right). \end{aligned}$$

Finally, for all $j \leq K$ making the change of variable $u_j = \frac{\left(u_j^H - u_j^L\right)(1+v_j)}{2}$, this is equivalent to

$$f_3\left[q, p, B(\eta), \theta\right] = \int_{-1}^{1} \cdots \int_{-1}^{1} \prod_{j=1}^{K} \left(\frac{u_j^H - u_j^L}{2}\right) dv. \tag{B.21}$$

Integrating out the random coefficients, the Case 3 likelihood is

$$\ln f_3\left(q, p, \theta\right) = \int_{\eta} f_3\left[q, p, B(\eta), \theta\right] \phi(\eta) d\eta.$$

C Data

Tables C.1-C.3 show the most frequently purchased two-item combinations. For completeness, Table C.1 is the same as Table A.3 in the paper.

The tables show the following. While each of the top 5 or 10 two-item combinations has an appreciable market share, in aggregate the top 5 account for only 54.34% of two-item combinations and the top 10 account for only 67.20%. To account for 95% of two-item combinations one must include 105 distinct combinations, which are all the combinations listed in Tables C.1-C.3 below. Most of these combinations have small market shares individually, but together they account for a large share of all two-item baskets. Our model can account for this wide variation in choices of types of fruit, numbers of types chosen, and the quantities of each.

Table C.1: Most frequently purchased 2-item combinations (A)

	Freq.	Pct.	Cum. Pct.
Banana, Apples	101533	25.03	25.03
Banana, Berries+Currants	52141	12.85	37.88
Banana, Easy Peelers	24442	6.03	43.91
Banana, Grapes	23977	5.91	49.82
Apples, Easy Peelers	18363	4.53	54.34
Berries+Currants, Apples	15931	3.93	58.27
Apples, Grapes	12052	2.97	61.24
Berries+Currants, Grapes	8592	2.12	63.36
Avocado, Banana	7915	1.95	65.31
Banana, Pears	7681	1.89	67.20
Apples, Pears	6299	1.55	68.76
Banana, Orange	5746	1.42	70.17
Berries+Currants, Easy Peelers	5506	1.36	71.53
Apples, Orange	5070	1.25	72.78
Easy Peelers, Grapes	4856	1.20	73.98
Banana, Melons	3551	0.88	74.85
Banana, Nectarines	3244	0.80	75.65
Banana, Lemon	3187	0.79	76.44
Banana, Kiwi Fruit	3144	0.78	77.21
Berries+Currants, Cherries	3018	0.74	77.96
Banana, Plums	2916	0.72	78.68
Avocado, Berries+Currants	2514	0.62	79.30
Banana, Cherries	2511	0.62	79.92
Berries+Currants, Melons	2151	0.53	80.45
Berries+Currants, Nectarines	2133	0.53	80.97
Apples, Kiwi Fruit	2043	0.50	81.48
Apples, Lemon	2009	0.50	81.97
Apples, Melons	1898	0.47	82.44
Banana, Grapefruit	1829	0.45	82.89
Apples, Nectarines	1803	0.44	83.33
Apples, Plums	1790	0.44	83.77
Avocado, Apples	1751	0.43	84.21
Grapes, Pears	1745	0.43	84.64
Easy Peelers, Pears	1734	0.43	85.06
Grapes, Orange	1508	0.37	85.44

Note: The table records the frequency with which various 2-item combinations were purchased.

Table C.2: Most frequently purchased 2-item combinations (B)

	Freq.	Pct.	Cum. Pct.
Berries+Currants, Kiwi Fruit	1485	0.37	85.80
Berries+Currants, Orange	1426	0.35	86.15
Banana, Pineapples	1392	0.34	86.50
Berries+Currants, Plums	1391	0.34	86.84
Berries+Currants, Lemon	1285	0.32	87.16
Berries+Currants, Pears	1275	0.31	87.47
Apricot, Banana	1263	0.31	87.78
Grapes, Kiwi Fruit	1262	0.31	88.09
Grapes, Melons	1237	0.30	88.40
Grapes, Plums	1201	0.30	88.69
Banana, Peaches	1126	0.28	88.97
Banana, Mango	1109	0.27	89.24
Easy Peelers, Plums	1087	0.27	89.51
Banana, Dates	1060	0.26	89.77
Easy Peelers, Orange	1060	0.26	90.04
Apples, Grapefruit	986	0.24	90.28
Grapes, Nectarines	980	0.24	90.52
Easy Peelers, Melons	963	0.24	90.76
Easy Peelers, Lemon	949	0.23	90.99
Berries+Currants, Pineapples	899	0.22	91.21
Grapes, Lemon	871	0.21	91.43
Berries+Currants, Peaches	870	0.21	91.64
Easy Peelers, Kiwi Fruit	861	0.21	91.85
Berries+Currants, Mango	842	0.21	92.06
Apples, Pineapples	818	0.20	92.26
Apples, Plums	1790	0.44	83.77
Avocado, Apples	1751	0.43	84.21
Grapes, Pears	1745	0.43	84.64
Easy Peelers, Pears	1734	0.43	85.06
Grapes, Orange	1508	0.37	85.44
Berries+Currants, Kiwi Fruit	1485	0.37	85.80
Berries+Currants, Orange	1426	0.35	86.15
Banana, Pineapples	1392	0.34	86.50
Berries+Currants, Plums	1391	0.34	86.84
Berries+Currants, Lemon	1285	0.32	87.16

Note: The table records the frequency with which various 2-item combinations were purchased.

Table C.3: Most frequently purchased 2-item combinations (C)

	Freq.	Pct.	Cum. Pct.
Berries+Currants, Pears	1275	0.31	87.47
Apricot, Banana	1263	0.31	87.78
Grapes, Kiwi Fruit	1262	0.31	88.09
Grapes, Melons	1237	0.30	88.40
Grapes, Plums	1201	0.30	88.69
Banana, Peaches	1126	0.28	88.97
Banana, Mango	1109	0.27	89.24
Easy Peelers, Plums	1087	0.27	89.51
Banana, Dates	1060	0.26	89.77
Easy Peelers, Orange	1060	0.26	90.04
Apples, Grapefruit	986	0.24	90.28
Grapes, Nectarines	980	0.24	90.52
Easy Peelers, Melons	963	0.24	90.76
Easy Peelers, Lemon	949	0.23	90.99
Berries+Currants, Pineapples	899	0.22	91.21
Grapes, Lemon	871	0.21	91.43
Berries+Currants, Peaches	870	0.21	91.64
Easy Peelers, Kiwi Fruit	861	0.21	91.85
Berries+Currants, Mango	842	0.21	92.06
Apples, Pineapples	818	0.20	92.26
Orange, Pears	818	0.20	92.47
Nectarines, Plums	791	0.19	92.66
Cherries, Apples	774	0.19	92.85
Lemon, Orange	741	0.18	93.03
Avocado, Easy Peelers	699	0.17	93.21
Easy Peelers, Nectarines	691	0.17	93.38
Apricot, Berries+Currants	673	0.17	93.54
Apples, Mango	664	0.16	93.71
Pears, Plums	618	0.15	93.86
Apples, Peaches	611	0.15	94.01
Avocado, Grapes	575	0.14	94.15
Grapes, Pineapples	572	0.14	94.29
Cherries, Grapes	556	0.14	94.43
Lemon, Lime	542	0.13	94.56
Grapes, Grapefruit	513	0.13	94.69

Note: The table records the frequency with which various 2-item combinations were purchased.

Another way to see the variety of choices and the potential role of complementarities is to look at the frequency of basket size conditional on fruit choice. Tables C.4-C.5 show, conditional on purchase of a fruit type, how frequently each basket size was purchased. Except for bananas, cherries, and lemons, all categories are more likely to be purchased in combinations than as stand-alone categories. The relative frequencies of basket size vary across fruit categories and the larger baskets are usually less frequent. These patterns strongly violate the usual independence assumptions of typical discrete choice demand models.

Table C.4: Number of categories purchased conditional on fruit type (A)

	Size of fruit basket						
	1	2	3	4	5	6	Total
Apricot	425	618	656	560	409	681	3349
	12.69	18.45	19.59	16.72	12.21	20.33	100.00
Avocado	5099	4592	3903	2879	1938	2399	20810
	24.50	22.07	18.76	13.83	9.31	11.53	100.00
Banana	121133	103981	71415	39854	20041	15468	371892
	32.57	27.96	19.20	10.72	5.39	4.16	100.00
Berries+Currants	46458	37782	28220	18430	11102	10739	152731
	30.42	24.74	18.48	12.07	7.27	7.03	100.00
Cherries	2611	3296	2778	2040	1336	1731	13792
	18.93	23.90	20.14	14.79	9.69	12.55	100.00
Dates	1104	867	703	494	285	416	3869
	28.53	22.41	18.17	12.77	7.37	10.75	100.00
Apples	59971	76517	59414	34882	18040	14545	263369
	22.77	29.05	22.56	13.24	6.85	5.52	100.00
Easy Peelers	30193	35914	30488	18977	10402	9099	135073
	22.35	26.59	22.57	14.05	7.70	6.74	100.00
Grapes	36085	39580	33187	22622	13088	11627	156189
	23.10	25.34	21.25	14.48	8.38	7.44	100.00
Grapefruit	2387	2985	2930	2567	1857	2522	15248
	15.65	19.58	19.22	16.83	12.18	16.54	100.00
Kiwi Fruit	4297	6561	6821	5705	4081	5062	32527
	13.21	20.17	20.97	17.54	12.55	15.56	100.00
Lemon	8175	7736	6671	5183	3601	4227	35593
	22.97	21.73	18.74	14.56	10.12	11.88	100.00
Lime	975	1372	1302	1082	835	1211	6777
	14.39	20.24	19.21	15.97	12.32	17.87	100.00
Lychees	182	210	226	170	126	200	1114
	16.34	18.85	20.29	15.26	11.31	17.95	100.00

Note: The table records the frequency of each fruit basket size conditional on purchasing the listed fruit category. Column 1 lists the fruit categories. The middle columns record the frequencies. The final column records the total number of observations of each type.

Table C.5: Number of categories purchased conditional on fruit type (B)

	Size of fruit basket						
	1	2	3	4	5	6	Total
Mango	2074	2865	3059	2533	1830	2735	15096
	13.74	18.98	20.26	16.78	12.12	18.12	100.00
Melons	7669	9212	8553	6539	4494	5378	41845
	18.33	22.01	20.44	15.63	10.74	12.85	100.00
Nectarines	6141	8720	8061	6114	4187	4731	37954
	16.18	22.98	21.24	16.11	11.03	12.47	100.00
Orange	12404	15247	13809	9562	5739	5838	62599
	19.82	24.36	22.06	15.28	9.17	9.33	100.00
Passion Fruit	218	317	283	246	200	328	1592
	13.69	19.91	17.78	15.45	12.56	20.60	100.00
Paw-Paws	138	219	234	216	154	261	1222
	11.29	17.92	19.15	17.68	12.60	21.36	100.00
Peaches	2811	3855	3528	2667	1766	2247	16874
	16.66	22.85	20.91	15.81	10.47	13.32	100.00
Pears	11486	20541	22356	16794	10240	9645	91062
	12.61	22.56	24.55	18.44	11.25	10.59	100.00
Pineapples	4857	5352	4905	3959	2734	3675	25482
	19.06	21.00	19.25	15.54	10.73	14.42	100.00
Plums	8947	11592	10874	8150	5423	5893	50879
	17.58	22.78	21.37	16.02	10.66	11.58	100.00
Pomegranates	559	565	454	346	262	288	2474
	22.59	22.84	18.35	13.99	10.59	11.64	100.00
Rhubarb	356	393	380	293	209	236	1867
	19.07	21.05	20.35	15.69	11.19	12.64	100.00
Sharon Fruit	341	375	371	340	266	366	2059
	16.56	18.21	18.02	16.51	12.92	17.78	100.00
Total	377096	401264	325581	213204	124645	121548	1563338
	24.12	25.67	20.83	13.64	7.97	7.77	100.00

Note: The table records the frequency of each fruit basket size conditional on purchasing the listed fruit category. Column 1 lists the fruit categories. The middle columns record the frequencies. The final column records the total number of observations of each type.

D Hedonic price functions

As discussed in Section 6.2 in the paper, for each fruit category we estimate a hedonic price model

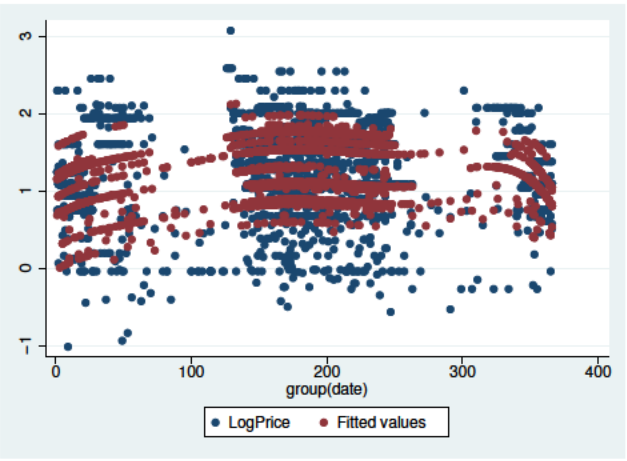
$$\ln p_{it} = \beta x_{it} + h(t) + \varepsilon_{it}$$

where $\ln p_{it}$ is the price of item i in period t, x_{it} is a vector of characteristics of item i in period t and $h(t)$ is a 6th order polynomial function of time. Time is measured as the day within the year. Characteristics included in the regressions are country of origin, branded, organic, tiering (economy, premium or standard), fascia (one of ten firms in the UK or other), packaging, online shop, and small store.

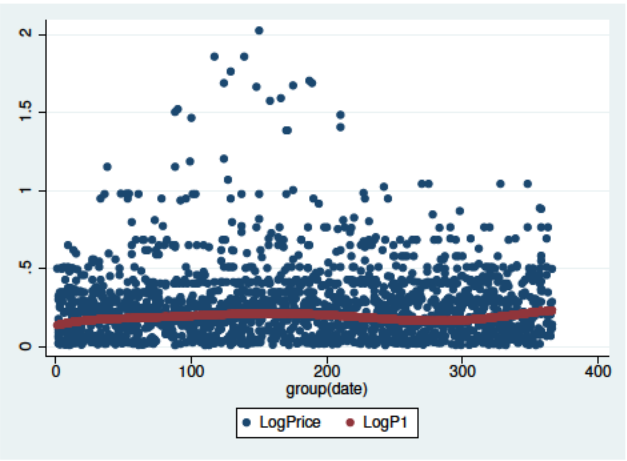
Figure D.1 shows price data and imputed prices for 3 representative examples of the 27 fruit categories: apricots, bananas and cherries. Price is observed for each shopping trip where a particular fruit is purchased. Each figure shows a scatter plot of observed log prices and imputed log prices. For apricots and cherries, prices rise in the spring and the autumn. These are periods when fresh apricots and cherries are more costly and more scarce. In contrast, the price of bananas is relatively flat. The pictures also make clear that at a single point in time there is a great deal of variability in price. This variation is primarily due to variation across fascia and variation due to promotions.

Figure D.1: Prices of apricots, bananas and cherries

(a) Apricots



(b) Bananas



(c) Cherries