

# Sparse demand systems: corners and complements

Arthur Lewbel

Boston College and IFS

and

Lars Nesheim

CeMMAP, IFS and UCL

First draft March 2013, revised October 2018

## **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 our model to the demand for fruit in the United Kingdom, analyzing 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 our 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. We apply the model to the demand for fruit in the United Kingdom (UK), and use the results to simulate the impacts of a variety of scenarios, including price changes that could be induced by Brexit, or by a change in the Value Added Tax (VAT), or by a proposed merger of two large supermarket chains. We find that fruit demand has multiple attributes that neither standard discrete choice nor standard continuous demand systems would adequately explain.

Our model nests both standard continuous demand systems (quadratic utility functions with Gorman (1976, 1980) and Lancaster (1966) type consumption technologies) and standard discrete choice models (multinomial probit models 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 to consume zero quantities of most types of goods, and includes substantial unobserved preference heterogeneity. Our model also does not impose a two stage structure where consumers must first choose which goods to buy and then how much of each good to buy, or vice versa.

As our motivating example, we consider consumer demand for fresh fruit in the UK, obtained from scanner data. In a typical store, there are more than a couple of dozen different 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 varies 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, the standard assumptions that underlie this methodology are likely to be seriously violated in many empirical applications, including our application to fruit demand. The assumptions underlying estimation methods using traditional continuous demand systems (such as those described in Deaton and Muellbauer 1980) are likewise violated when many goods are bought in zero quantities.

Methods to overcome the violated assumptions in either the discrete or continuous demand models are either nonexistent or intractable in applications like ours. 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. Alternatively, considering the problem as a traditional continuous demand model will also be intractable, because most consumers will be at corner solutions (buying zero of most goods) most of the time. In contrast, our model is tractable for such problems and captures consumer substitution patterns at both the intensive (continuous choice) and extensive (discrete choice) margins.

Due to nonuniform size and the ability to purchase multiple items, fruit can be purchased in close to continuous quantities. For this issue, extensions to multinomial choice models exist that combine a discrete choice for a product with a continuous choice for the quantity of that product, going back to Dubin and McFadden (1984).

A far more serious limitation of multinomial choice models for our application is that they generally rule out complements. Complementarities are important in a wide range of empirical applications. For example, some fruits are strong complements for others (e.g., different types of berries are frequently purchased together 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. As noted above, our data yields 80,730 distinct combinations to be considered. And this number would skyrocket higher if we even crudely discretized the quantities purchased of each fruit. We could alternatively allow for some complementarities 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, particularly when considering both the types of fruit chosen and the quantity of each.

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 or implicit price models; 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 separate impacts on each type of fruit.

Aggregate demand elasticities for individual types of fruit are determined by the demand responses of individual consumers at both the intensive and the extensive margins. They also depend on the sum of responses across many disparate households. An analysis that aggregates fruit purchases cannot capture the separate impacts of these two margins nor can it capture the heterogeneity in individual responses. In response to an increase in the price of apples, some people might buy fewer apples, some might switch from buying apples to pears, and still others might drop apples entirely and start buying peaches and nectarines. Because different households buy different baskets of fruit, individual elasticities vary significantly across households. Our results discussed in Section 7.1 illustrate these types of effects at the household level. The estimated elasticities reported in Section 7.3 show the net effect of these responses on aggregate elasticities, obtained by summing the intensive and extensive effects across many different disparate consumers. These results have important implications for welfare calculations, construction of price indices, market structure, and tax policies.

The basic structure of our model incorporates a Gorman (1976, 1980) and Lancaster (1966) type characteristics model into a continuous demand system with substantial unobserved preference heterogeneity. The model then allows for many corner solutions in the demand for 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. With a limiting version of the matrix  $B$ , our model also 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.

Then 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, and some Appendices provide additional technical material and estimation details.

## 2 Literature

As summarized by, e.g., Blundell and Meghir (1987), the literature on continuous demand considers three main causes that might generate observations with zero purchases of some goods. One cause 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. More simply, this occurs when an individual never chooses to buy the good in question, regardless of income or relative price. An example would be the zero cigarette consumption of a nonsmoker. A second cause is infrequency of purchase. A good that is durable or storable may be consumed regularly, but purchases may be infrequently observed. A third cause is corner solutions. These occur when the price of a good is above its reservation levels so that nonnegativity constraints are binding. In such a case, at the given level of prices and total expenditures, a consumer chooses to purchase zero units of the good in question. Starting from such a corner solution, if the price of the good falls or if incomes or other goods' prices rise so that the price falls below the reservation price, then the consumer will switch to purchase a positive amount.

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.

Lexicographic preferences are typically modelled analogously to Heckman (1979) type sample selection models, where 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, specifically for demand systems, Yen and Lin (2006). A recent example of demand system estimation of this type (still with a small number of goods) is the semiparametric estimator of Sam and Zheng (2010). Models like these that have equations for determining zero versus non-zero consumption that are separate from equations determining desired quantities 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 under any circumstance. In addition, our model allows for substantial heterogeneity of preferences, and so accommodate the types of behaviour that lexicographic preferences seek to capture by having the value of parameters that would result in positive purchases of a good be set equal to extreme values for consumers who would never purchase the good. Our model allows for the possibility that consumers may purchase all goods but predicts that purchase of some baskets by some households are extremely rare.

An extreme version 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 maximize a profit function by choosing a single brand (of computer) along with deciding how many units to buy (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 Yorokoglu (2012) and Thomassen, Smith, Schiraldi, and Seiler (2017). In the latter paper the chosen good is a store, and the quantity is amount spent. This paper has some similarities to ours in that they allow trips to more than one store (analogous to buying more than one good), and they employ a quadratic utility function. Crucially, they do not exploit corners to generate zeros as we do, and so they require two stage budgeting, and cannot tractably handle more than two stores (i.e., two nonzero goods).

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 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 two flaws. First, 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 (AID), where the vector  $e$  appears in the demand functions  $f_j^*(P, Y, e)$  only inside a general price index, which is replaced by an approximate Stone price index. This is the model proposed by Heien and Wessells (1990).

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

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 is numerically intensive and 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.

In our model, zeros are handled using both corners and the Gorman (1976, 1980) and Lancaster (1966) characteristics model of taste heterogeneity. Dubois, Griffith, and Nevo (2014) also exploit a Gorman Lancaster characteristics model, but 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).

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  $j$  that is 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 or consumers, but for now, omit that to ease notation. Suppose that the utility a consumer gets from purchasing this bundle of goods is a function of  $K$  unobserved 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 our 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 all 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. Still, 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,  $K$ .

However, an additional feature of our model is that we allow that maximized utility may have many corners, i.e., points where indifference curves intersect with axes in attribute space. This results in utility being maximized with anywhere from zero to  $K$  different types of goods. The more corners, 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 then results not only in different consumers purchasing different baskets of goods, but it also results in different consumers facing different corners, and different numbers of corners. The result is that our model can encompass the variation seen in real data, where consumers vary in the numbers of goods that they buy (from zero to  $K$ ), and vary in the choice of which goods to buy, and vary in the quantities they purchase of each nonzero good.

Some other models exist that deal with zeros in an analogous way, such as the censored Almost Ideal Demand System discussed earlier, but these models have the drawbacks described in the previous section (one of the most serious drawbacks being that the errors in those models do not formally correspond to utility parameters).

Assume that each individual chooses  $q$  to maximize the utility function

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

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 corresponding to satiation levels or bliss points for each attribute  $k$ .

This utility function is quadratic and concave in  $q$  which allows us to employ standard quadratic programming techniques to handle zeros coming from corners in our model. These methods are computationally fast and efficient 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. Our theory would largely go through with more general utility functions that are concave in  $e - Bq$ , but would be computationally far more burdensome.

This utility function nests both standard continuous demand systems and standard discrete choice models. Continuous quadratic direct utility functions correspond to the special case of this model with  $K = J$  and with  $B^T e \geq p$ , ensuring a full dimensional interior solution. 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. With a limiting version of the matrix  $B$ , we show that our model also nests standard multinomial logit or

probit models as a special case. 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 most of our derivations 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 <sup>1</sup>, and substituting the budget constraint into the objective, the consumer chooses  $q$  to maximize

$$y - p^T q - 0.5 (e - Bq)^T (e - Bq) \text{ such that } q \geq 0. \quad (1)$$

### 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  $\delta$  is a vector of Lagrange multipliers. The first order conditions are

$$\begin{aligned} 0 &= -p - B^T (Bq - e) + \delta \\ 0 &= \delta^T q, \quad \delta \geq 0, \quad q \geq 0. \end{aligned} \quad (2)$$

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  $E \subset \mathbb{R}^K$  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 convex 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.

---

<sup>1</sup>The utility function in (1) can be multiplied by any positive number without changing any predictions or implications of the model.

### 3.2 Piecewise linear demands

To prove identification, it is useful to characterize solutions that have the maximum number  $K$  of nonzero elements. To do so, let  $\bar{q} = (\bar{q}_1, \bar{q}_2)$  be a vector for which  $\bar{q}_1 > 0$  and  $\bar{q}_2 = 0$  such that  $\dim(\bar{q}_1) = K$ . Without loss of generality, the elements of  $\bar{q}_1$  can be taken to be the first  $K$  elements of  $\bar{q}$ . Let  $p_1$  and  $p_2$  be the corresponding price subvectors and  $B_1$  and  $B_2$  the corresponding submatrices of  $B$  so that

$$B = \begin{bmatrix} B_1 & B_2 \end{bmatrix}.$$

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  $\bar{q}$  is optimal for all  $e$  satisfying

$$-p_1 - B_1^T (B_1 \bar{q}_1 - e) = 0 \quad (3)$$

$$-p_2 - B_2^T (B_1 \bar{q}_1 - e) \leq 0 \quad (4)$$

$$\bar{q}_1 \geq 0 \quad (5)$$

Equation (3) defines the inverse demand curve  $-p_1 = B_1^T (B_1 \bar{q}_1 - e)$ . The inequalities (4) and (5) define conditions under which choosing  $\bar{q}_1 > 0$  and  $\bar{q}_2 = 0$  is optimal. When  $B_1$  is nonsingular the system can be simplified and solved to provide explicit conditions describing the piecewise linear demand function

$$\bar{q}_1 = (B_1^T B_1)^{-1} (B_1^T e - p_1) \quad (6)$$

$$p_2 - B_2^T (B_1^T)^{-1} p_1 \geq 0 \quad (7)$$

$$\bar{q}_1 \geq 0. \quad (8)$$

In words, the nonnegative  $\bar{q}_1$  is optimal if it satisfies the demand equation (6) and if the projection of its price vector is cheaper than the price vector  $p_2$ . When (4) is not binding, small changes in  $p_2$  have no impact on demand for  $\bar{q}_1$ .

## 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

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

where  $\delta$  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 (6) becomes

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

For empirical application, one could then let  $B$  or  $e$  depend on product characteristics  $z$ , consumer characteristics  $x$ , or unobserved heterogeneity  $\eta$  as detailed in Section 5.2 below. For example, for each household  $h$ , 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 that could be readily estimated.

## 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 (1) reduces to

$$\max \{y - p^T q - 0.5 (Bq - e_1)^2\} \quad (9)$$

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

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

which is maximized either at an interior point of the feasible range of values of  $q_j$  given by the first order condition  $-p_j - (b_{1j} q_j - e_1) b_{1j} = 0$  or at one of the endpoints of the feasible range, i.e., either  $q_j = 0$  or  $q_j = y/p_j$ . Since fruit is a small component of total expenditures, we assume that  $y$  is large enough to make the solution  $q_j = y/p_j$  (spending all of one's money on a single type of fruit) suboptimal. At the interior point, the optimal quantity is  $q_j = (e_1 - p_j/b_{1j})/b_{1j}$  (which is only feasible if  $e_1 > p_j/b_{1j}$ ) which yields utility  $y + 0.5 ((p_j/b_{1j}) - e_1)^2 - 0.5 e_1^2$ . Otherwise, at the corner solution, if  $q = 0$  is optimal, then utility is  $y - 0.5 e_1^2$ .

It follows that if  $e_1 \leq \min_{\ell \in \{1, \dots, J\}} \{p_\ell/b_{1\ell}\}$  for all goods  $j$ , then utility is maximized by  $q = 0$ . Otherwise, utility is maximized by buying the quantity  $q_j = (e_1 - p_j/b_{1j})/b_{1j}$  of the good  $j = \arg \min_{\ell \in \{1, \dots, J\}} \{p_\ell/b_{1\ell}\}$ , and not buying any other good. Let  $j = 0$  denote not buying any of the goods (i.e., choosing  $j = 0$  corresponds to choosing  $q = 0$ ). Let  $p_0 = 1$  and  $\ln b_{10} = -\ln e_1$ . It then follows that equation (9) implies making a discrete choice to purchase good  $j = \arg \max_{\ell \in \{0, 1, \dots, J\}} \{\ln b_{1\ell} - \ln p_\ell\}$  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  $\ln b_{1jh} = (\beta_0 + \beta_1 x_h) z_j + \varepsilon_{jh}$  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

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

(which is multinomial probit if  $\varepsilon$  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

$$(y - p^T q) \beta_0 - 0.5 q^T B^T B q + u^T q \quad (10)$$

for some positive scalar  $\beta_0$  (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 (10) equal to equation (1) up to an affine transformation (multiplication by  $\beta_0$  and addition of  $e^T e$ ) that has no effect on consumer choices. Essentially, our original utility model can be derived from equation (10) by completing the square.

Now consider the limiting case of (10) in which  $B \rightarrow 0$  and  $Be \rightarrow u$  where  $u_j = \beta z_j + \varepsilon_j$ . In the limit case, maximizing (10) is equivalent to maximizing

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

With the addition of the constraint that  $\sum_j q_j \leq 1$ , this is the standard multinomial choice model. The parameters  $(\beta, \beta_0)$ , may depend on both observable demographics  $x_h$  and random coefficients  $\eta_h$ . In particular, if one assumes that  $\varepsilon_j$  are distributed as independent Type I extreme value random variables, then this is the standard multinomial logit model that is estimated in transport economics, labour, marketing, and industrial organization.

## 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  $\{e : B^T e \geq 0\}$ . 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 (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 (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  $\tilde{B}_j = [B_j \ B_{-j}]$  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 the those  $K$  goods. Identification of the  $j$ 'th column of  $B$  is assured using expressions like (7) and (6) with nonsingularity of  $\tilde{B}_j$  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  $E = \{e \mid B^T e \geq 0\}$ .  $e$  is distributed independently of  $p$ .*

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  $r \in \{1, \dots, R\}$  index each possible element of  $\mathcal{B}$ . Let  $i^r = \{i_1^r, \dots, i_k^r\}$  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  $\tilde{R} \subset \mathcal{B}$  denote the smallest set of bundles such that  $B^r$  is nonsingular for all  $r \in \tilde{R}$  and  $B_j$  is a column in  $B^r$  for some  $r$ . The set  $\tilde{R}$  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  $r \in \tilde{R}$ , there is a set  $A \subseteq P \times Y$  and a set  $Q^r = \{q_r \in \mathbb{R}^J \text{ with } q_r(i) = 0 \text{ if } i \notin i^r\}$  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 (7) and (6) are satisfied for  $\bar{q} = q_r$ . That is when  $p_{-r} - B_{-r}^T (B_r^T)^{-1} p_r \geq 0$  and  $q_r = (B_r^T B_r)^{-1} (B_r^T e - p_r) \geq 0$ . 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  $p' = (p'_r, p'_{-r})$  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  $\bar{q}_r$  denote  $K$  vectors of prices and quantities of those  $K$  goods. By the first order conditions, a consumer buying  $\bar{q}_r$  has  $B_r^T B_r \bar{q}_r = B_r^T e - p_r$ . 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  $\bar{q}_r = (B_r^T B_r)^{-1} (B_r^T e - p_r)$ . 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  $\bar{q}_r$  conditioning on  $p$  (which can be calculated at any point that is not on the boundary of the region) is  $(B_r^T B_r)^{-1}$ , 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  $B_r^T e = B_r^T B_r \bar{q}_r + p_r$  for all observable  $\bar{q}_r, p_r$  pairs. Since  $\bar{q}_r$  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 externalities 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. Prices alone do not explain the consumption patterns in the data. Consumers facing the same prices choose different fruit baskets.

Let  $B^h$  be the matrix of preference parameters for household  $h$ , let  $B_j^h$  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  $\|B_j^h\| = \sqrt{\sum_{k=1}^K (b_{hkj})^2}$ , plays a similar role. When  $\|B_j^h\|$  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. 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  $(d_{hj}, C_j^h) = H(B_j^h)$  where  $d_{hj} = \|B_j^h\|$ ,  $\|C_j^h\| = 1$ , and  $H$  is the hyperspherical coordinate transformation detailed in Appendix A. 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

$$\ln d_{hj} = z_j^T \beta_h \quad (11)$$

$$c_{hkj} = \pi \Phi^{-1}(z_{kj}^T \gamma_h) \quad \forall k \leq \min(K, j) - 2 \quad (12)$$

$$c_{hkj} = 2\pi \Phi^{-1}(z_{kj}^T \gamma_h) \quad \forall k = \min(K, j) - 1 \quad (13)$$



where  $\Phi$  is the normal CDF,  $(z_j, z_{kj})$  are vectors of product characteristics, and  $(\beta_h, \gamma_h)$  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.

We assume

$$\beta_{hj} = \beta_{j0} + \beta_{j1}^T x_h + \beta_{j2}^T \eta_h \quad (14)$$

$$\gamma_{hkj} = \gamma_{kj0} + \gamma_{kj1}^T x_h + \gamma_{kj2}^T \eta_h \quad (15)$$

where  $x_h$  is a vector of observable demographic variables and  $\eta_h$  is a  $N_\eta$  dimensional vector of low dimensional unobservable latent factors. The  $J \times N_\eta$  matrix  $\beta_2 = [\beta_{12}, \dots, \beta_{j2}, \dots, \beta_{J2}]$  is an upper triangular matrix of factor loadings mapping the low dimensional  $\eta$  into the random coefficients  $\beta_h$ . 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  $\gamma_2 = [\gamma_{122}, \dots, \gamma_{KJ-1,2}]$  is an upper triangular matrix of factor loadings mapping the low dimensional  $\eta$  into the random coefficients  $c_h$ . Note that the matrices of factor loadings are normalized to be upper triangular.

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  $\eta$ . In our empirical application we assume that the dimension of  $\eta_h$  is either 2 or 3 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 (11) and (12).

### 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  $\bar{q}_r = (B_r^T B_r)^{-1} (B_r^T e - p_r)$ . Conditioning on this price regime, the conditional distribution of  $\bar{q}_r$  given prices nonparametrically identifies the distribution of  $(B_r^T B_r)^{-1} (B_r^T e - p_r)$  given  $p_r$ .

The parameterized distributions of  $B$  and  $e$  are then identified as long as their parameters can be recovered from moments of  $(B_r^T B_r)^{-1}$  and of  $(B_r^T B_r)^{-1} B_r^T e$ . 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  $\eta_h$  is a  $N_\eta$  dimensional vector of independent normally distributed latent factors. We then estimate the factor loadings  $\beta_{j2}$  and  $\gamma_{kj2}$ . 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  $(p_h, q_h, x_h)$  for each household  $h$ . Income  $y_h$  plays no role due to the quasilinear utility assumption. We assume that  $e \sim N(\mu, \Sigma)$ , that  $\eta \sim N(0, I)$ , and that  $\dim(\eta) = 2$ . The parameters of the model are  $\theta = (\mu, \Sigma, \beta, \gamma)$  where  $\beta = \{\beta_{j0}, \beta_{j1}, \beta_{j2}\}_{j=1}^J$  is the vector of all parameters in (11) and  $\gamma = \{\gamma_{kj0}, \gamma_{kj1}, \gamma_{kj2}\}$  is the vector of all parameters in (12). The parameters  $(\mu, \Sigma)$  determine the distribution of  $e$ . They primarily determine the number of items chosen and the quantities purchased. The parameters  $(\beta, \gamma)$  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.

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 as detailed in Appendix B 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.

Full details of the likelihood function are given in Appendix B.

## 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 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 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.

Tables 3-5 show 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 3-5. 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 6-7 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 very 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 of 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 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.

### 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 and hence observed prices are either zero or extremely low. Fruits (such as ugly 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, as can be seen in Figure 1, 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 is 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,

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), and  $N_\eta$ , the dimension of unobserved latent factors. For  $J = 27$ ,  $K = 5$ , and  $N_\eta = 2$ , there are 321 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. 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 what the model predicts for individual consumers, we plot predicted demand for 9 household types. Each type is defined by a realization of the vector  $\eta$  such that for each type each element of  $\eta$  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, Figure 2 shows the frequency of purchased basket sizes when prices vary from 50% to 200% of baseline prices (basket size here refers to the number of different types of fruit, not the quantities of each). These household types rarely purchase baskets with 5 items. The modal choice is 2 items for households 1, 3 and 6, 3 items for households 2, 4, 7, and 9 and 4 items for households 5 and 8. For all households, the number of items purchased varies from 2 to 4 or 5 except for household 6 who almost always buys 2 items.

For each household type, we also examined how the basket composition varies with price. For these 9 household types, all baskets contain bananas. Apples are purchased by all households except household 2 and pears are purchased by all except 1 and 4. Demand for the other fruits varies across households. Some households bought easy peelers (households 1, 2, 4, and 8) and some bought nectarines (households 1, 3, 5, and 6). Only households 2 and 4 bought grapes and only household 1 bought peaches.

Finally, Figures 3 and 4 plots these households' demand curves for various fruits, as functions of the price of bananas and of the price of apples, respectively. Turning first to Figure 3, the demand curve for bananas is downward sloping in the price of bananas for all households, but with varying slopes, and these demand curves are kinked for all households other than household 9. The effects of banana prices on other fruits varies widely; some cross price effects are negative, some are positive, some are flat, and some switch signs depending on which combination of products are purchased. Figure 4 shows a similar wide range of effects on fruit demand by household type as a function of apple prices.

## 7.2 Aggregate demand predictions

Figures 5 - 7 show estimated aggregate demand curves for each fruit. They also show what fraction of 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 demand for bananas shown in Figure 5 panel (c), has a relatively gentle slope, the demand for avocados in panel (b) is steeper and has more curvature, and the demand for apricots in panel (a) is very steep for low prices before flattening substantially for prices above about £3.4 per kg. These graphs also show how frequently each fruit appears in baskets of various sizes. For example, most purchases of bananas are in baskets with three or fewer types of fruits, while dates, which are rarely consumed, are usually purchased in a basket with four different types of fruit.

## 7.3 Elasticities

Tables 8 - 11 show estimates of average own and cross price elasticities. All but 4 of the own price elasticities are between -1.14 (easy peelers) and -7.54 (lychees). Three of the products have own price elasticities less than one in magnitude: bananas (-0.709), berries (-0.870), and apples (-0.598). These are three of the top 5 products by market share. This suggests that these products might at least sometimes be sold as loss leaders, on sale for a relatively low

price, despite inelastic demand. One of fruit own price elasticities is very large in magnitude; apricots (-14.7). This is not altogether unexpected. Apricots are purchased in only 0.21 % of transactions. While they are not the smallest market share products, 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 curves for apricots seen in Figure 6 panel (a) shows that the demand curve for apricots is very steep when prices are low and have flatter demand curves as prices rise.

Typical discrete choice models assume all goods are substitutes and so do not permit zero or positive 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 8 - 11 show that about a third of all pairs of fruits have zero cross price effects.

Among the nonzero cross-price elasticities there is 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 Table 8, when the price of bananas rises, demand for apricots cherries and dates rise suggesting that, on average they are substitutes. In contrast, demand for apples go down slightly suggesting that they are complements. The table also show that demand for bananas is relatively insensitive to the price of other fruits, while other fruits are more sensitive to the price of bananas. Note also that all of these elasticities are averages across consumers. Because we have random coefficients, it is possible for some pairs of fruits to be substitutes for some households and complements for others.

## 7.4 Counterfactual scenarios

Many current large scale shifts in the economy could affect the markets for fruit in the UK. For example, Brexit may 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 is that two of the largest supermarket chains, Asda and Sainsbury's, who account for 16.8% and 15.5% of UK fruit sales respectively, have proposed a merger. Such a merger could increase their market power, possibly driving up fruit prices.

At the same time, various tax policy changes could be under consideration 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  $\tau_{1j} = 0.1s_{EU,j}$  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  $\tau_{2j} = 0.1s_{UK,j}$  where  $s_{UK,j}$  is the share of fruit sourced from the UK. For scenario 3, the percentage price increase is assumed to be  $\tau_{3j} = 0.05(s_{ASDA,j} + s_{SAIN,j})$  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 13. 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 has large impacts on the prices of 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 has moderate impacts on prices because only berries, cherries, apples, pears and rhubarb have significant UK supplies. 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 12 and 14.

The second column in Table 12 shows that the EU tariff has a small percentage impact on most fruit categories but a big negative impact on apricots, kiwis, nectarines, peaches and pomegranates. For all of these categories except pomegranates, the tariff leads to a more than 5% price increase. The results show that the impacts are quite large, larger than one would predict based on the own price elasticities alone. This suggests that the cross price effects are quite important. Also, the table shows that the tariff increases demand for several fruits that are primarily sourced from outside the EU. These fruits (lychees, mangos, passion fruits, and paw-paws) have low levels of demand to begin with. Their prices are unaffected by the tariff and yet their demand increases substantially.

Scenarios 2 and 3 have much more moderate impacts on prices and also on the resulting demand for fruit. In scenario 2, demand for cherries, dates and rhubarb fall 7.37%, 5.05%, and 56.6% respectively. All other responses are less than 5%. In scenario 3, only apricots (-9.06%) and rhubarb (-7.59%) experience changes in demand larger than 5%.

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 8 categories, 10-20%



for 12 categories and more than 20% for 5 categories. The latter scenario reduces demand by less than 10% for 3 categories, 10-20% for 6 categories and more than 20% for 18 categories.

Table 14 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 0.415% to 2.02%. In all cases, other than the UK cost shock scenario, the effect is increasing in the percentile of expenditure. The percentage impacts are two to three times larger for high spending households than for low spending households.

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 12.5 pence per shopping trip and costs the 90th percentile households about 40 pence per shopping trip. The UK cost shock has very small impacts, less than 5 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, other than the merger, 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 12.6 pence per household per shopping trip, representing about a 2% decrease in revenues. The merger increases firm revenues by about 5 pence per trip and reduces consumer surplus by 11.1 pence per trip.

## 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.

## A Hyperspherical representation of $B$

Since  $B$  is upper triangular,  $b_{kj} = 0$  if  $k > j$ . The number of nonzero elements in column  $B_j$  is  $\bar{k} = \min\{K, j\}$ . Let  $C_j = [c_{1j}, \dots, c_{\bar{k}-1}]^T$ . 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

$$\begin{aligned} B(1, j) &= d_j \cos(c_{1j}) \\ B(2, j) &= d_j \sin(c_{1j}) \cos(c_{2j}) \\ B(3, j) &= d_j \sin(c_{1j}) \sin(c_{2j}) \cos(c_{3j}) \\ &\vdots \\ B(\bar{k}-1, j) &= d_j \sin c_{1j} \cdots \sin(c_{\bar{k}-2}) \cos(c_{\bar{k}-1}) \\ B(\bar{k}, j) &= d_j \sin(c_{1j}) \cdots \sin(c_{\bar{k}-2}) \sin(c_{\bar{k}-1}) \end{aligned} \tag{16}$$

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

## B Estimation details

The likelihood function for the model is defined 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

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. Inverting the demand function given in Section 3.2 equation (6), inverse demand is

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

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

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

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

## 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

$$-p_1 - B_1^T (B_1 q_1 - e) = 0 \quad (17)$$

$$-p_2 - B_2^T (B_1 q_1 - e) \leq 0 \quad (18)$$

$$q_1 > 0. \quad (19)$$

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  $\tilde{e} = U^T e$  and partition  $\tilde{e} = (\tilde{e}_1, \tilde{e}_2)$  where  $\tilde{e}_1$  is ( $d_1 \times 1$ ) and  $\tilde{e}_2$  is ( $d_2 \times 1$ ). Then rewrite (17) as

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

or

$$VS_1 \tilde{e}_1 = p_1 + B_1^T B_1 q_1. \quad (20)$$

For each  $q_1$  there are multiple vectors  $\tilde{e}$  that solve (20). In fact, there is a linear space of dimension  $d_2$ . In other words, for each  $(q_1, \tilde{e}_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ , there is a unique  $\tilde{e}_1$  defined by

$$\tilde{e}_1 = G_0 p_1 + G_1 q_1 \quad (21)$$

where

$$\begin{aligned} G_0 &= S_1^{-1} V^T \\ G_1 &= S_1^{-1} V^T (B_1^T B_1). \end{aligned} \quad (22)$$

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

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

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

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

where  $(G_0, G_1)$  are defined in (22).

We observe  $q_1$  if inequality (18) is satisfied. Since  $B_1 = USV^T$  and  $e = U\tilde{e}$ , this is equivalent to

$$-p_2 - B_2^T U (SV^T q_1 - \tilde{e}) \leq 0. \quad (23)$$

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

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

where  $\tilde{B}_{21}$  is size  $(d_1 \times J - d_1)$  and  $\tilde{B}_{22}$  is size  $(d_2 \times J - d_1)$ , inequality (23) is

$$-p_2 - \begin{bmatrix} \tilde{B}_{21}^T & \tilde{B}_{22}^T \end{bmatrix} \left( \begin{bmatrix} S_1 V^T q_1 \\ 0 \end{bmatrix} - \begin{bmatrix} \tilde{e}_1 \\ \tilde{e}_2 \end{bmatrix} \right) \leq 0$$

or

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

Substituting from equation (21) this is equivalent to

$$\tilde{B}_{22}^T \tilde{e}_2 \leq p_2 - \tilde{B}_{21}^T G_0 p_1 + \tilde{B}_{21}^T (S_1 V^T - G_1) q_1. \quad (24)$$

Rewrite (24) as

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

where

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

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

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

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

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

$$f_2 [q, p, B(\eta), \theta] = \int f_{q_1 \tilde{e}_2} (q_1, \tilde{e}_2) 1 (M_1 \tilde{e}_2 \leq M_2) d\tilde{e}_2. \quad (25)$$

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

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

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

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

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

Therefore, (25) can be written

$$f_2 [q, p, B(\eta), \theta] = \int f_{q_1 z_2} (q_1, z_2) 1 \left( \widetilde{M}_1 z_2 \leq \widetilde{M}_2 \right) dz_2 \quad (26)$$

where

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

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

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

$$f_2 [q, p, B(\eta), \theta] = \int_{RQz_2 \leq D} \widetilde{f}_{q_1 z_2} (q_1, z_2) \frac{e^{-0.5 z_2^T z_2}}{(2\pi)^{\frac{d_2}{2}}} dz_2 \quad (27)$$

$$= \int_{Rx \leq D} \widetilde{f}_{q_1 z_2} (q_1, Q^{-1}x) \frac{e^{-0.5 x^T x}}{(2\pi)^{\frac{d_2}{2}}} dx \quad (28)$$

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  $j \in J_{d_2}^+$ ,

$$-\infty \leq x_{d_2} \leq \frac{D_j - \sum_{i < d_2} R(j, i) x_i}{R(j, d_2)}$$

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

$$\frac{D_j - \sum_{i < d_2} R(j, i) x_i}{R(j, d_2)} \leq x_{d_2} \leq \infty.$$

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

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

and

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

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

$$f_2[q, p, B(\eta), \theta] = \int_{x_1^L}^{x_1^H} \cdots \int_{x_{d_2}^L}^{x_{d_2}^H} \tilde{f}_{q_1 z_2}(q_1, Q^{-1}x) \frac{e^{-0.5x^T x}}{(2\pi)^{\frac{d_2}{2}}} dx. \quad (29)$$

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

$$f_2[q, p, B(\eta), \theta] = \int_{u_1^L}^{u_1^H} \cdots \int_{u_{d_2}^L}^{u_{d_2}^H} \tilde{f}_{q_1 z_2}(q_1, Q^{-1}x(u)) du \quad (30)$$

where

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

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

$$f_2[q, p, B(\eta), \theta] = \int_{-1}^1 \cdots \int_{-1}^1 \prod_{j=1}^{d_2} \left( \frac{u_j^H - u_j^L}{2} \right) \tilde{f}_{q_1 z_2}(q_1, Q^{-1}x(v)) dv. \quad (31)$$

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

The conditional density function  $f_2$  depends on the parameters  $\theta$  and on the random coefficient  $\eta$ . Integrating out the random coefficients, the Case 2 likelihood function is

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

### B.3 Case 3: Choice of 0 goods

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

$$-p + B^T e \leq 0. \quad (32)$$

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

$$B^T e \leq p. \quad (33)$$

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

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

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

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

$$\tilde{B} = QR$$

be the QR decomposition of  $\tilde{B}$  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

$$f_3 [q, p, B(\eta), \theta] = \int_{R^T x \leq p - B^T \mu} \frac{e^{-0.5x^T x}}{(2\pi)^{\frac{K}{2}}} dx. \quad (34)$$

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  $x_K \in [x_K^L, x_K^H]$  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

$$f_3 [q, p, B(\eta), \theta] = \int_{x_1^L}^{x_1^H} \cdots \int_{x_K^L}^{x_K^H} \frac{e^{-0.5x^T x}}{(2\pi)^{\frac{d_2}{2}}} dx. \quad (35)$$

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

$$f_3[q, p, B(\eta), \theta] = \int_{u_1^L}^{u_1^H} \cdots \int_{u_K^L}^{u_K^H} du \quad (36)$$

where

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

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

$$f_3[q, p, B(\eta), \theta] = \int_{-1}^1 \cdots \int_{-1}^1 \prod_{j=1}^K \left( \frac{u_j^H - u_j^L}{2} \right) dv. \quad (37)$$

Integrating out the random coefficients, the Case 3 likelihood is

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

## C References

- Amano, N. (2018), "Nutrition Inequality: The Role of Prices, Income, and Preferences," Yale University 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.
- 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.
- 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.
- Lancaster, K. (1966), "A New Approach to Consumer Theory", *Journal of Political Economy*, 74, 132-157.
- Lee, L., and M. M. Pitt, (1986), "Microeconomic 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.
- 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 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.
- 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.

## D Tables

Table 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 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 3: 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 4: 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 5: 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.

Table 6: Number of categories purchased conditional on fruit type (A)

	Size of fruit basket						Total
	1	2	3	4	5	6	
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 7: Number of categories purchased conditional on fruit type (B)

	Size of fruit basket						Total
	1	2	3	4	5	6	
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.

Table 8: Elasticities (1)

Price	Apricots	Avocados	Bananas	Berries	Cherries	Dates	Apples
$p_{Apricots}$	<b>-14.7</b>	0	0.00781	-0.000131	0	0	-0.00123
$p_{Avocados}$	-7.35e-12	<b>-1.93</b>	-0.00451	-0.0232	0.226	0.00919	-0.0107
$p_{Bananas}$	2.45	-0.0626	<b>-0.709</b>	0.0832	0.155	0.331	-0.103
$p_{Berries}$	-0.059	-0.464	0.12	<b>-0.87</b>	-0.261	-0.972	0.135
$p_{Cherries}$	0	0.204	0.0101	-0.0117	<b>-2.63</b>	0.0252	-0.0174
$p_{Dates}$	0	0.00221	0.00573	-0.0117	0.00672	<b>-3.37</b>	0
$p_{Apples}$	-0.878	-0.338	-0.236	0.214	-0.613	0	<b>-0.598</b>
$p_{EasyPeelers}$	3.36	0.677	0.212	0.00107	-0.354	0.175	0.0895
$p_{Grapes}$	-0.0136	-0.359	0.0439	0.488	-0.0112	0.00153	-0.031
$p_{Grapefruits}$	0	-0.000842	0.000163	-0.0014	0.000223	0.236	-0.000314
$p_{Kiwis}$	-0.306	0.134	0.0221	-0.0123	0	0	0.00365
$p_{Lemons}$	-0.0446	-0.0398	0.0131	0.0572	0.44	0.411	-0.0197
$p_{Limes}$	0	0.00484	0.0078	-0.000514	0.0275	0.178	-8.7e-05
$p_{Lychees}$	0	0	-0.000343	0.000538	0	0.0499	0
$p_{Mangos}$	0	0.111	-0.0121	-0.013	0.222	0	0.00117
$p_{Melons}$	-0.041	0.0126	-0.056	-0.0931	0.00674	0.223	-0.0312
$p_{Nectarines}$	6.82	0.0501	-0.0452	-0.00887	0.275	0	0.0232
$p_{Oranges}$	-0.0139	0.253	-0.0122	-0.0495	0.273	1.49	-0.0233
$p_{Passionfruits}$	0	0.0237	0.00295	0	0.0193	0	-0.000288
$p_{Paw-paws}$	0	0.00193	0.000361	-1.15e-05	0.0263	0	-0.000176
$p_{Peaches}$	0	0.00303	-0.00464	0.0299	0.303	-0.000169	0.0458
$p_{Pears}$	0	0.394	0.197	-0.0322	0.525	0.00124	0.0297
$p_{Pineapples}$	0	-0.0135	0.0192	0.0194	-0.0327	0	2.47e-06
$p_{Plums}$	0.551	0.0122	-0.00602	-0.0885	0.0354	-0.0113	0.125
$p_{Pomegranates}$	-0.532	0.00265	0.00256	0.000682	0	0	0.000675
$p_{Rhubarb}$	0	0.0109	-0.0011	0.000239	0	0	-4.63e-05
$p_{Sharonfruits}$	0	-0.00701	0.0053	0.0123	0.00277	0.492	-0.00207

Table 9: Elasticities (2)

Price	Easy Peelers	Grapes	Grapefruits	Kiwis	Lemons	Limes	Lychees
$p_{Apricots}$	0.00548	-2.78e-05	0	-0.00297	-0.00066	0	0
$p_{Avocados}$	0.0249	-0.0166	-0.00315	0.0293	-0.0133	0.0197	0
$p_{Bananas}$	0.108	0.028	0.00846	0.0674	0.0607	0.44	-0.384
$p_{Berries}$	0.000786	0.452	-0.105	-0.0539	0.383	-0.0418	0.866
$p_{Cherries}$	-0.0117	-0.000465	0.000754	0	0.132	0.101	0
$p_{Dates}$	0.00155	1.7e-05	0.213	0	0.0331	0.174	0.967
$p_{Apples}$	0.104	-0.0454	-0.0372	0.0254	-0.208	-0.0112	0
$p_{EasyPeelers}$	<b>-1.14</b>	-0.176	0.178	0.372	-0.102	-0.121	-0.0614
$p_{Grapes}$	-0.14	<b>-1.28</b>	0.175	-0.0458	0.535	-0.00518	0
$p_{Grapefruits}$	0.00175	0.00215	<b>-3.48</b>	0	0.0323	0.0194	0.243
$p_{Kiwis}$	0.0624	-0.00964	0	<b>-2.39</b>	-0.0579	0	0
$p_{Lemons}$	-0.0112	0.074	0.362	-0.038	<b>-2.48</b>	0.518	2.43
$p_{Limes}$	-0.00109	-5.88e-05	0.0179	0	0.0426	<b>-5.32</b>	0.239
$p_{Lychees}$	-2.81e-05	0	0.0113	0	0.0101	0.0121	<b>-7.54</b>
$p_{Mangos}$	-0.00935	0.0166	0	0	0.0373	0.0579	0
$p_{Melons}$	-0.0741	0.118	1.26	0.000817	0.129	1.06	0
$p_{Nectarines}$	0.00835	-0.0304	0	0.000693	0.0514	0	0
$p_{Oranges}$	-0.0834	-0.0434	0	0.00133	0.0053	0.529	0
$p_{Passionfruits}$	-0.00154	0.00244	0	0.00158	0	0	0
$p_{Paw-paws}$	7.16e-05	-2.44e-05	0	-0.000206	0.000829	0	0
$p_{Peaches}$	0.0346	0.116	-0.00455	-0.0334	0.123	0	0
$p_{Pears}$	0.369	-0.168	0	0.0602	-0.0267	0.0658	0
$p_{Pineapples}$	0.000512	0.0392	0	0.0266	0.00675	0	0
$p_{Plums}$	0.112	0.00717	-0.00239	0.83	-0.135	0	0
$p_{Pomegranates}$	0.00105	4.11e-05	0	0.00203	0.0081	0	0
$p_{Rhubarb}$	-0.000242	0.00033	0.0107	0	0.00705	0	0
$p_{Sharonfruits}$	-0.0258	0.0153	0.256	0	0.205	0.656	0

Table 10: Elasticities (3)

Price	Mangos	Melons	Nectarines	Oranges	Passion fruits	Paw-paws	Peaches
$p_{Apricots}$	0	-0.000205	0.155	-7.8e-05	0	0	0
$p_{Avocados}$	0.149	0.00142	0.0258	0.0321	0.399	0.349	0.001
$p_{Bananas}$	-0.224	-0.0876	-0.323	-0.0215	0.688	0.908	-0.0213
$p_{Berries}$	-0.349	-0.21	-0.0912	-0.125	-2e-11	-0.0415	0.198
$p_{Cherries}$	0.267	0.000685	0.128	0.0312	0.293	4.29	0.0904
$p_{Dates}$	0	0.00606	0	0.0455	0	0	-1.35e-05
$p_{Apples}$	0.0496	-0.111	0.378	-0.0936	-0.153	-1.01	0.481
$p_{EasyPeelers}$	-0.34	-0.227	0.117	-0.287	-0.703	0.352	0.311
$p_{Grapes}$	0.482	0.287	-0.338	-0.119	0.889	-0.0955	0.834
$p_{Grapefruits}$	0	0.0379	0	0	0	0	-0.000403
$p_{Kiwis}$	0	0.00042	0.00163	0.000768	0.121	-0.17	-0.0505
$p_{Lemons}$	0.149	0.0436	0.0791	0.00201	0	0.449	0.122
$p_{Limes}$	0.019	0.0293	0	0.0165	0	0	0
$p_{Lychees}$	0	0	0	0	0	0	0
$p_{Mangos}$	<b>-4.7</b>	0.0456	0.253	0.024	0	0	0.0299
$p_{Melons}$	0.541	<b>-1.81</b>	0.0612	0.401	0.0461	0	0.0297
$p_{Nectarines}$	0.659	0.0134	<b>-2.67</b>	0.0483	0	0	0.0833
$p_{Oranges}$	0.253	0.357	0.197	<b>-1.83</b>	0.547	0	0.212
$p_{Passionfruits}$	0	0.000309	0	0.00412	<b>-3.12</b>	0	-0.0128
$p_{Paw-paws}$	0	0	0	0	0	<b>-6.63</b>	2.43e-05
$p_{Peaches}$	0.12	0.0101	0.129	0.0809	-0.65	0.0133	<b>-3.64</b>
$p_{Pears}$	0.558	0.279	0.1	0.434	0	0.738	0.0703
$p_{Pineapples}$	0.0734	0	0	0	0	0	0.0347
$p_{Plums}$	0.0124	-0.000398	0.0075	-0.00211	-0.262	-0.439	-0.0724
$p_{Pomegranates}$	0	0	-0.00361	0.00295	0.00605	0	0
$p_{Rhubarb}$	0	0.00498	0	0.00733	0	0	-2.64e-05
$p_{Sharonfruits}$	-0.167	0.0244	0.0196	0.0823	0	0	0

Table 11: Elasticities (4)

Price	Pears	Pineapples	Plums	Pomegranates	Rhubarb	Sharon fruits
$p_{Apricots}$	0	0	0.00329	-0.2	0	0
$p_{Avocados}$	0.0285	-0.0068	0.00165	0.0226	0.167	-0.0101
$p_{Bananas}$	0.198	0.135	-0.0112	0.303	-0.233	0.106
$p_{Berries}$	-0.0466	0.196	-0.239	0.116	0.0735	0.353
$p_{Cherries}$	0.0342	-0.0149	0.00431	0	0	0.00359
$p_{Dates}$	2.17e-05	0	-0.000368	0	0	0.17
$p_{Apples}$	0.0682	3.96e-05	0.533	0.182	-0.0225	-0.0944
$p_{EasyPeelers}$	0.726	0.00703	0.41	0.244	-0.101	-1.01
$p_{Grapes}$	-0.263	0.429	0.021	0.00759	0.11	0.477
$p_{Grapefruits}$	0	0	-8.6e-05	0	0.0437	0.0984
$p_{Kiwis}$	0.0199	0.0612	0.51	0.0788	0	0
$p_{Lemons}$	-0.00579	0.0102	-0.0543	0.206	0.324	0.881
$p_{Limes}$	0.00117	0	0	0	0	0.232
$p_{Lychees}$	0	0	0	0	0	0
$p_{Mangos}$	0.0302	0.0277	0.00125	0	0	-0.18
$p_{Melons}$	0.179	0	-0.000476	0	0.678	0.311
$p_{Nectarines}$	0.0141	0	0.00197	-0.0598	0	0.0548
$p_{Oranges}$	0.248	0	-0.00225	0.199	0.889	0.934
$p_{Passionfruits}$	0	0	-0.0021	0.00307	0	0
$p_{Paw-paws}$	0.000295	0	-0.000327	0	0	0
$p_{Peaches}$	0.0155	0.053	-0.0294	0	-0.00122	0
$p_{Pears}$	<b>-2.15</b>	0.000314	-0.0263	0.0819	0	1.76
$p_{Pineapples}$	4.49e-05	<b>-2.35</b>	0.0793	0	0	0
$p_{Plums}$	-0.0141	0.298	<b>-1.97</b>	-0.522	0	0
$p_{Pomegranates}$	0.000696	0	-0.00827	<b>-3.85</b>	0	0
$p_{Rhubarb}$	0	0	0	0	<b>-6.03</b>	0
$p_{Sharonfruits}$	0.0885	0	0	0	0	<b>-5.36</b>

Table 12: 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	3.44	6.52%	-7.15e-11%	2.21%	-10%	20%
Avocados	4.3	1.93%	6.19e-11%	1.73%	-10%	20%
Bananas	2.55	-4.67e-11%	-4.67e-11%	1.56%	-10%	20%
Berries	7.49	3.54%	4.3%	1.75%	-10%	20%
Cherries	7.1	3.47%	1.9%	1.45%	-10%	20%
Dates	1.54	0.05%	1.97e-11%	1.08%	-10%	20%
Apples	4.91	2.86%	1.54%	1.66%	-10%	20%
Easy Peelers	5.28	4.24%	-6.45e-11%	1.71%	-10%	20%
Grapes	5.84	3.39%	0.08%	1.46%	-10%	20%
Grapefruits	2.39	0.87%	-3.6e-11%	1.75%	-10%	20%
Kiwis	4.84	6.65%	7.56e-11%	1.55%	-10%	20%
Lemons	4.41	5.21%	-1.67e-11%	1.7%	-10%	20%
Limes	2.89	2.27%	-7.48e-11%	1.91%	-10%	20%
Lychees	1.37	-6.87e-11%	-6.87e-11%	1.4%	-10%	20%
Mangos	3.87	6.5e-11%	6.5e-11%	1.38%	-10%	20%
Melons	2.6	1.81%	-8.08e-11%	1.51%	-10%	20%
Nectarines	4.17	5.96%	9.81e-11%	1.91%	-10%	20%
Oranges	4.94	2.88%	8.79e-11%	1.66%	-10%	20%
Passion fruits	1.8	6.02e-12%	6.02e-12%	1.51%	-10%	20%
Paw-paws	1.22	-1.02e-10%	-1.02e-10%	1.48%	-10%	20%
Peaches	2.57	8.97%	-1.67e-11%	1.54%	-10%	20%
Pears	4.98	5.15%	1.54%	1.59%	-10%	20%
Pineapples	2.48	6.22e-11%	6.22e-11%	1.22%	-10%	20%
Plums	4.65	3.19%	0.26%	1.61%	-10%	20%
Pomegranates	3.27	3.71%	-5.3e-11%	1.14%	-10%	20%
Rhubarb	2.51	0.11%	9.87%	1.78%	-10%	20%
Sharon fruits	8.27	0.5%	9.95e-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 13: Percentage change in demand due to tax/price change

Fruit	Baseline (kg)	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5
		EU tariff	UK cost shock	Merger	Subsidy	VAT
Apricots	2.39e-08	-21.5%	-1.15%	-9.06%	29.6%	-29.7%
Avocados	4.32e-07	1.42%	-1.26%	-2.44%	14.3%	-22.7%
Bananas	1.01e-05	1.97%	0.566%	-0.644%	4.28%	-8.1%
Berries	4.96e-06	-1.3%	-3.75%	-0.692%	3.15%	-5.95%
Cherries	2.36e-07	-2.45%	-7.37%	-1.4%	14.9%	-23.4%
Dates	2.89e-07	5.69%	-5.05%	0.481%	7.42%	-14.6%
Apples	1.2e-05	0.0861%	-0.283%	-0.602%	3.97%	-7.31%
Easy Peelers	9.56e-06	-2.46%	0.805%	-1.18%	6.92%	-12.4%
Grapes	6.89e-06	-3.07%	1.83%	-1.24%	9.3%	-16.8%
Grapefruits	2.08e-07	3.49%	-0.241%	-2.51%	11.5%	-20.8%
Kiwis	1.75e-06	-11.5%	-0.0034%	-1.69%	12.3%	-19.6%
Lemons	1.26e-06	-9.77%	1.76%	-2.28%	13.5%	-20.4%
Limes	1.58e-07	-6.59%	0.357%	-4.62%	21.2%	-33.5%
Lychees	1.69e-08	11.8%	3.13%	-3.68%	36.9%	-52.4%
Mangos	3.59e-07	13.6%	-0.474%	-2.7%	26.5%	-40.4%
Melons	6.35e-06	-0.612%	-0.7%	-1.95%	13.9%	-23.2%
Nectarines	8.65e-07	-12%	1.35%	-3.95%	19.5%	-30.9%
Oranges	2.97e-06	-3.25%	0.245%	-2.2%	13%	-23.2%
Passion fruits	6.14e-08	7.31%	2.62%	-2.87%	18.9%	-28.1%
Paw-paws	8.42e-09	21%	23%	-1.85%	16.7%	-15.6%
Peaches	2.18e-06	-26.6%	2.37%	-1.81%	13.9%	-22.6%
Pears	5.15e-06	-7.35%	-3.61%	-1.14%	8.41%	-16%
Pineapples	1.48e-06	4.91%	1.34%	-0.829%	12.7%	-20.6%
Plums	2.97e-06	-1.43%	-0.786%	-1.24%	8.2%	-13.9%
Pomegranates	6.66e-08	-12.2%	0.677%	-3.41%	39.2%	-44.4%
Rhubarb	4.82e-08	8.78%	-56.6%	-7.49%	56.6%	-59.8%
Sharon fruits	1.56e-07	15.4%	2.51%	-2.54%	12.7%	-22.6%

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 14: 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	6.19	-0.369%	0.48%	0.0333%	0.225%	-2.94%
25th percentile	10	0.0411%	0.646%	0.224%	-1.65%	0.419%
50th percentile	15.3	0.559%	0.59%	0.377%	-3.19%	3.58%
75th percentile	21.8	0.892%	0.537%	0.503%	-3.68%	5.16%
90th percentile	29.7	0.841%	0.478%	0.494%	-3.58%	4.93%
Change in consumer surplus (GBP)						
10th percentile	6.24	-0.259	-0.0723	-0.126	0.855	-1.47
25th percentile	12.6	-0.366	-0.106	-0.184	1.28	-2.2
50th percentile	22.6	-0.529	-0.148	-0.265	1.78	-3.06
75th percentile	37.7	-0.73	-0.2	-0.35	2.31	-4.14
90th percentile	55.6	-0.87	-0.277	-0.425	2.73	-4.82
Per capita effects						
Consumer surplus (GBP)	27.5	-0.553	-0.157	-0.272	1.75	-3.14
Tax revenue (GBP)	0.0	1.02e-05	2.93e-06	5.1e-06	-3.43e-05	5.5e-05
Firm Revenue	0.000318	-7.97e-06	-1.26e-06	-3.72e-06	2.45e-05	-4.31e-05

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.



## E Figures

Figure 1: Prices of apricots, bananas and cherries

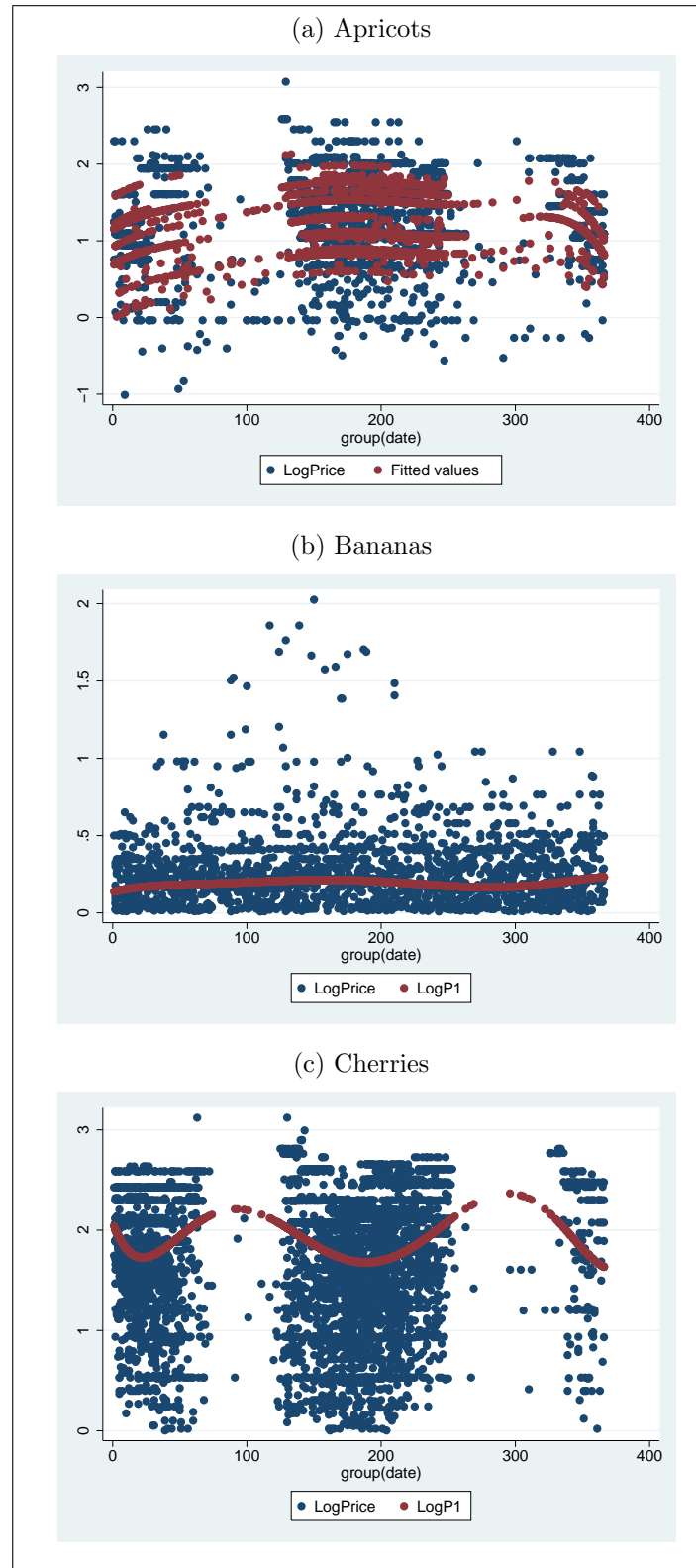
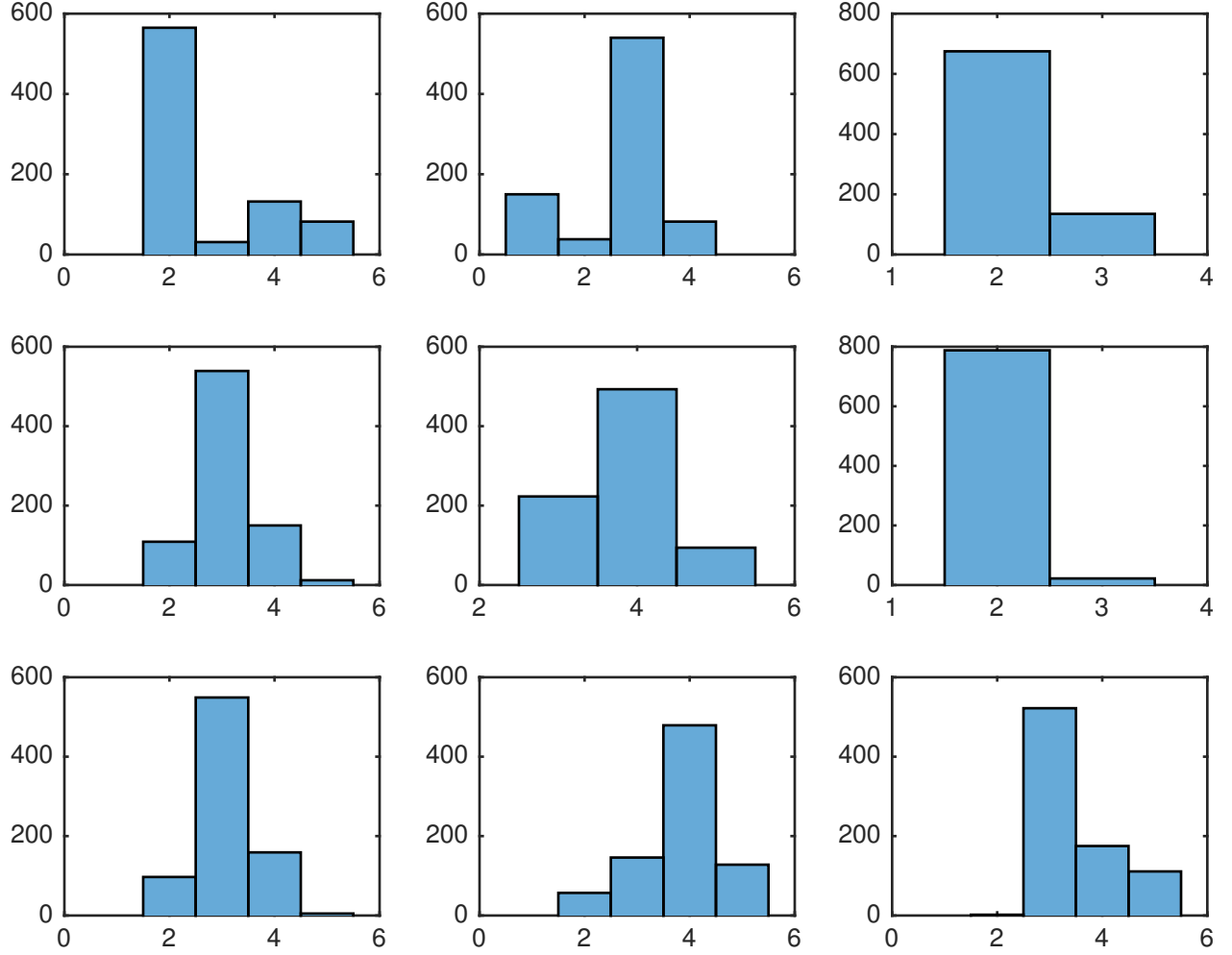
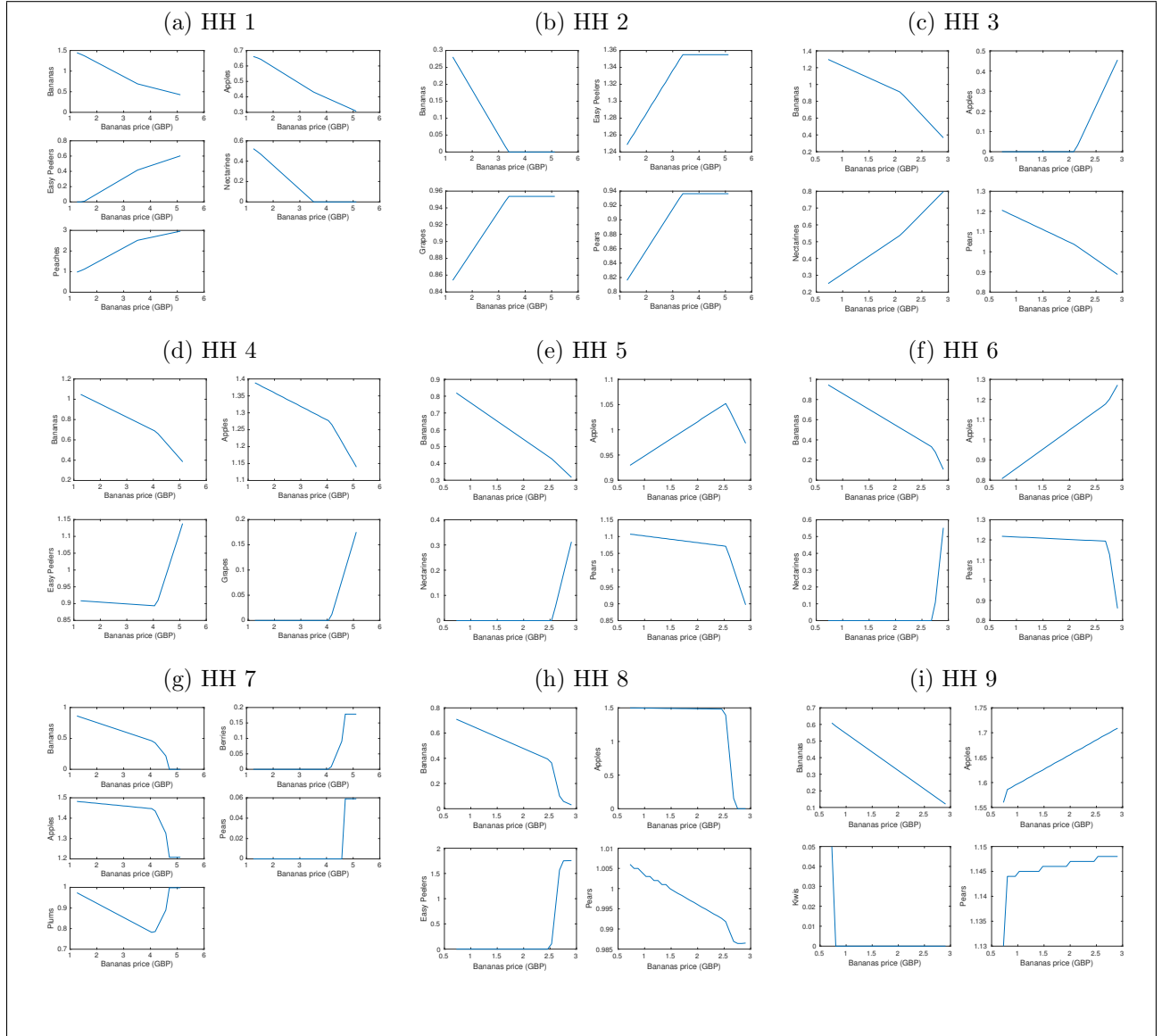


Figure 2: Frequency of basket size by household type



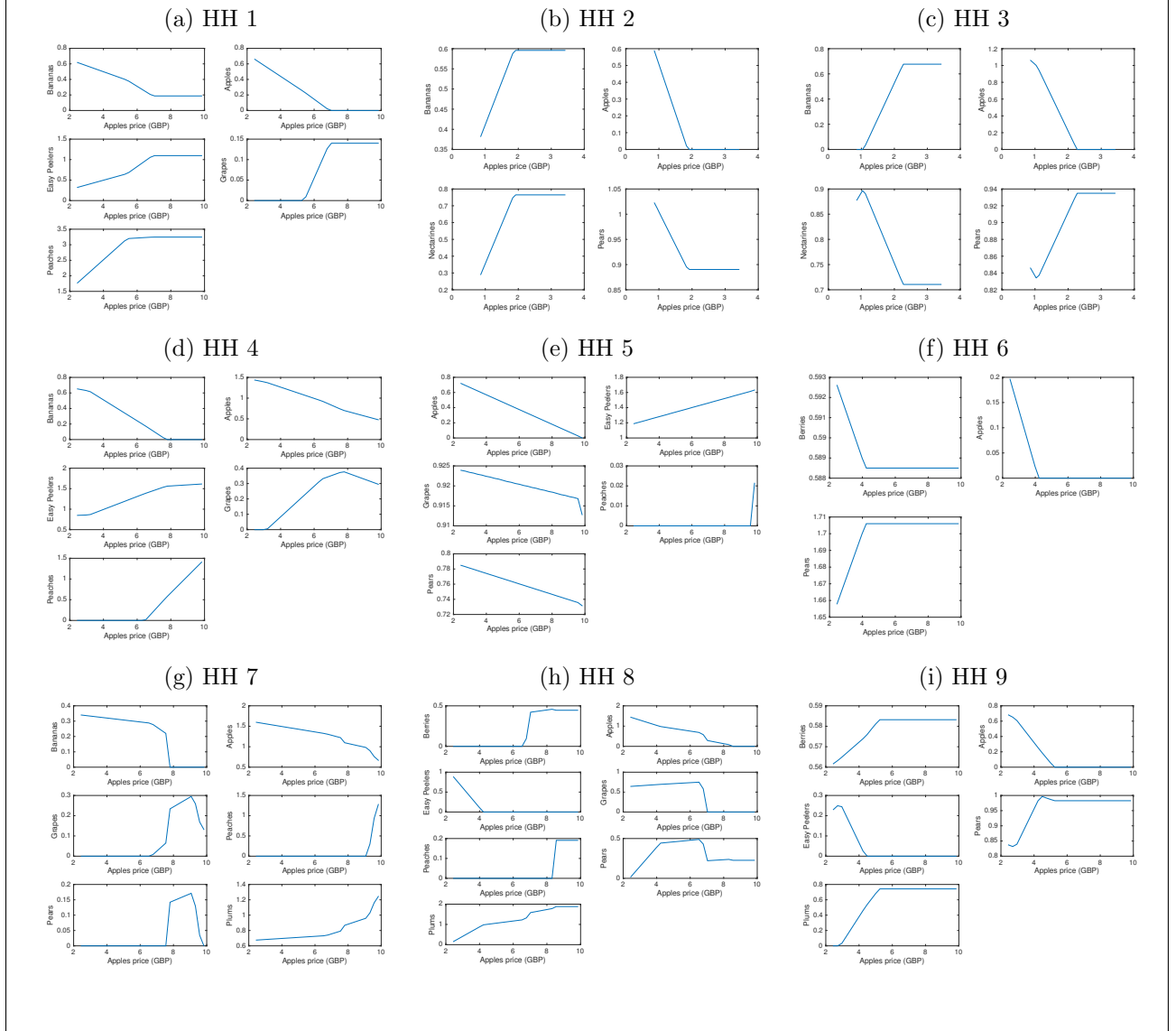
As prices vary from 50% - 200% of baseline, each panel shows the frequency of basket sizes for a single household type. Household types are defined by realizations of  $(\eta_1, \eta_2)$ . Each element of  $\eta_1$  and  $\eta_2$  takes one of three values drawn from the 25th, 50th, and 75th percentiles of the respective marginal distribution. Types 1-3 are in the top level panels, types 4-6 are in the middle three panels, and types 7-9 are in the bottom. Within a row, households are arranged in increasing order from left to right.

Figure 3: Demand vs. banana price: by household type



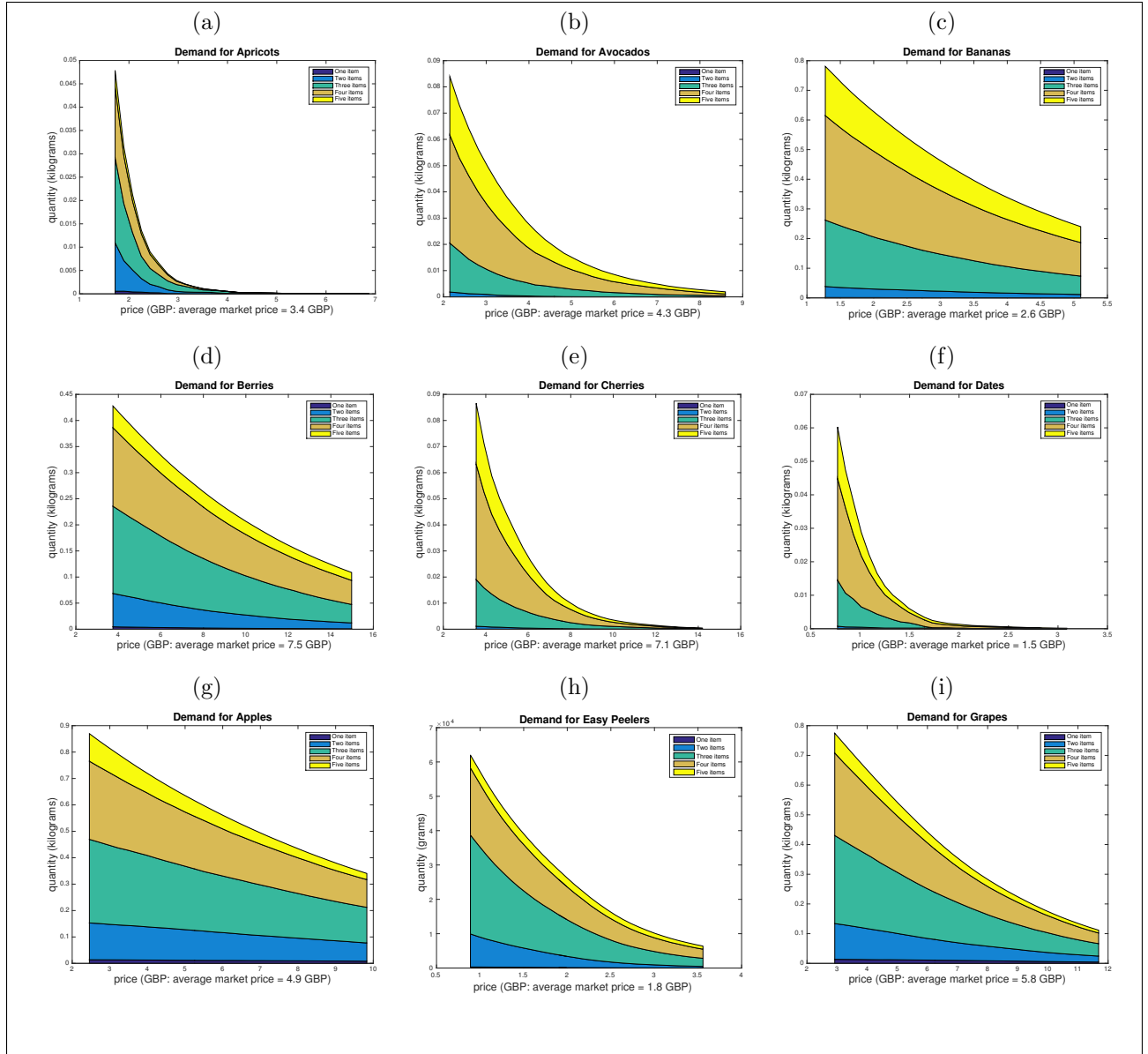
Each panel shows demand for fruit vs. banana price for one household type. Only fruits with non-zero demand are shown.

Figure 4: Demand vs. apple price: by household type



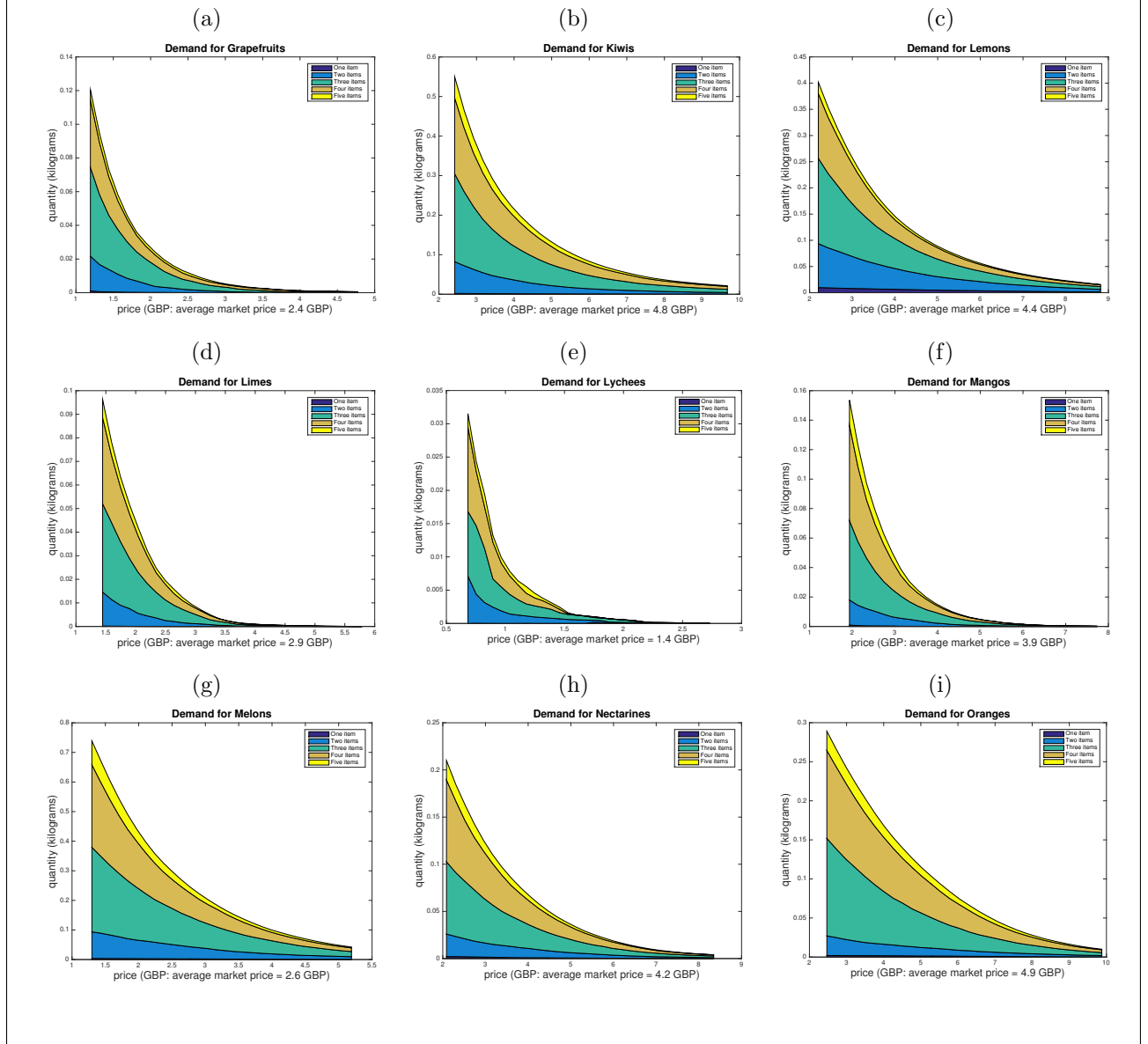
Each panel shows demand for fruit vs. apple price for one household type. Only fruits with non-zero demand are shown.

Figure 5: Aggregate demand curves (1)



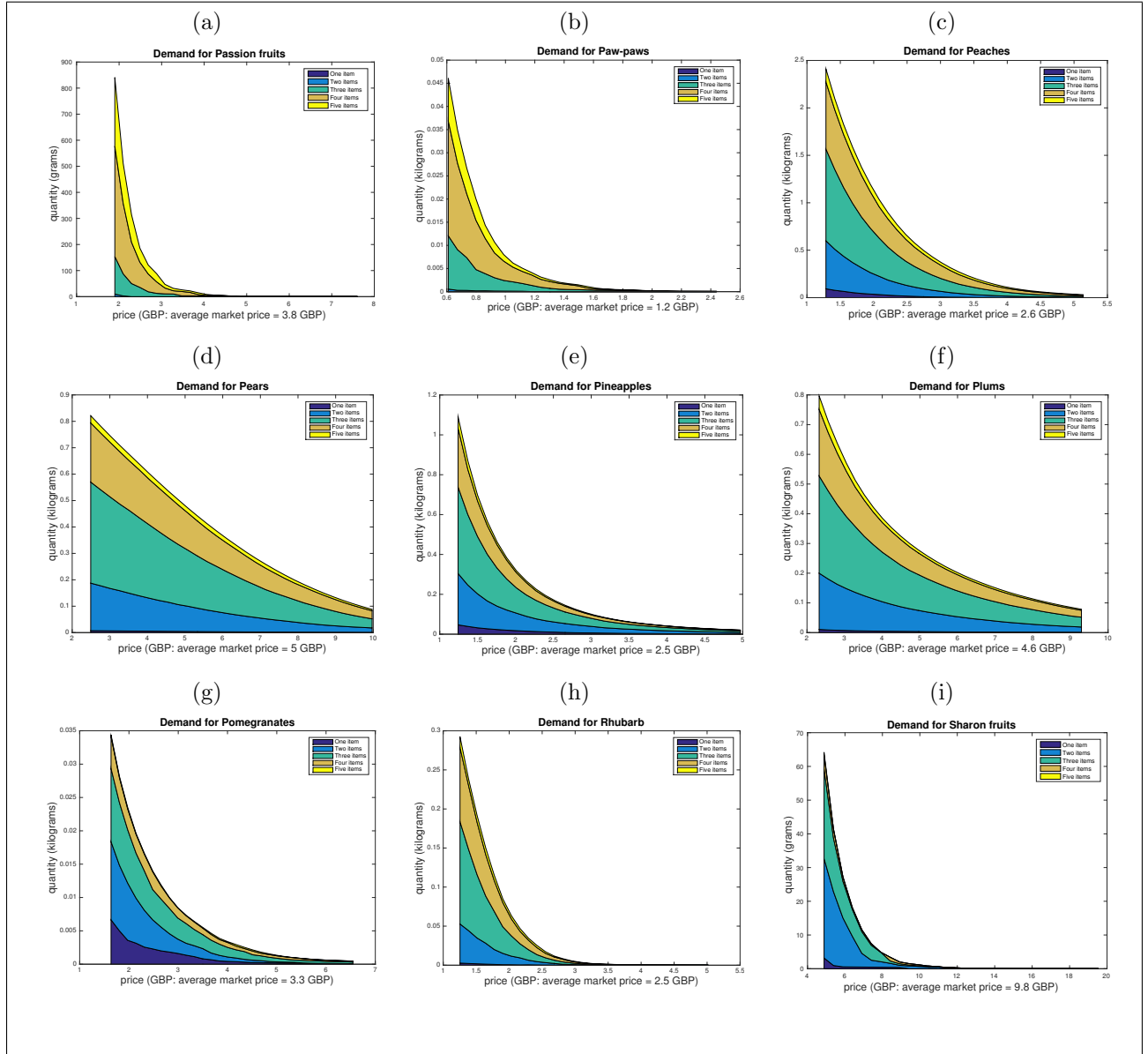
Each panel shows aggregate demand for a fruit category as that fruit's price varies from 50% to 200% of its baseline level. The figure also shows what fraction of demand comes from baskets of size 1 through 5.

Figure 6: 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 7: 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.