Consumer Choice over Shopping Bundles: A Linear Demand Approach

Afonso Rodrigues*

March 19, 2025

Abstract: It is common in retail settings to observe the joint purchasing of goods—*bundling* - as well as a vast number of products available for purchase. In this paper I show how ignoring bundling, a deviation from standard decision theory, may result in oversupply, mis-pricing, and inaccurate predictions of stock-out risks. I demonstrate that linear demand and quasi-linear quadratic utility models remain robust in this context, enabling a comprehensive theoretical characterisation of consumer choice over shopping bundles. The model is then empirically applied to estimate demand across several Portuguese supermarkets, illustrating how frequency of joint purchase may serve simultaneously as an effective proxy for approximating cross-price effects between goods.

1 Introduction

In economics, cross-price effects capture how the price change of one good influences the demand for another, reflecting the interconnected nature of consumer preferences across products. Capturing these cross-price effects accurately is essential for reliably modeling consumer choice, forecasting market outcomes, and guiding firm strategy. Retailers, particularly in markets characterised by extensive product offerings, face unique challenges in predicting these relationships, especially when managing inventory risks associated with product stock-outs; while existing demand models often excel in scenarios with a limited number of choices, they frequently encounter challenges when applied to the vast product portfolios typical of modern retail environments. This limitation becomes more acute in retail environments where joint purchasing of multiple goods — known as bundling — is commonplace,

*Department of Economics, University of Oxford (afonso.rodrigues@economics.ox.ac.uk). I thank Alexei Parakhonyak, Howard Smith, Ian Crawford, Alex Teytelboym, Abi Adams, Giulio Gottardo, and Leonardo Madio for their helpful insights, as well as the participants of the NIE PhD Symposium 2024 at the University of Warwick.

leading to further inter-dependencies among goods across categories.

In this paper, I show how quasi-linear quadratic utility (QQUM), and linear demand can be leveraged to provide unique insights on consumer choice over large choice sets, in markets where joint purchase of goods, *bundling*, is the norm.

Consumers may choose over small numbers of bundles, or be influenced by only a few latent factors (like common needs, seasonality, or promotions), rather than evaluating each product independently. I reveal how these deviations from standard decision theory may cause dependency in joint consideration frequency across goods, and establish an equivalence between consumer choice over bundles and that over goods given a dimensionality constraint. To verify the impact of model mis-specification, first, I conduct comparative statics over a Monte Carlo simulation with both imperfectly complementary and substitute goods. Mis-specification of consumer choice is found to on average lead to oversupply of goods. For a differentiated market of single-product firms, mis-pricing leads to on average 59.4% higher production levels than the profit-maximising level; 37% higher consumer surplus; but, surprisingly, a 16.4% aggregate profit gain. For a single monopolist retailer in the same scenario, the effects are more muted. Pricing over incorrect forecasts matches one of the profit-maximising equilibria, meaning that only the forecasts themselves raise concerns; these are on average off by 5%. Notably, production is highest in the correctly-specified monopolist multi-product retailer, followed by that amongst the single-product firms in the incorrect specification, and only then that amongst those in the correct specification. This result is key in demonstrating that the effects of bundling are qualitatively different from those of product complementarity. Complementary goods, when sold by the same firm, are sold more cheaply and at a higher rate. Where goods' purchases are linearly dependent, however, the incentive is to sell more expensive and in lower quantities. This result holds across all simulations. Whether mis-pricing in fact raises aggregate profits depends on the

product mix.

Second, I analyse the implications of forecasting errors for measuring the impact of product stock-outs. In general, these errors lead to (i) overstatement of consumer surplus and aggregate profit losses, (ii) compounding risks due to mis-pricing, and (iii) over-demand relative to forecasts for a significant number of goods, despite average oversupply. Firm size heterogeneity is somewhat less apparent. For single-product firms (supermarkets), the impact of product stock-outs on consumer surplus is roughly 73.7% (83.9%) less than forecast, though this rises to 81.6% (83.7%) when looking at aggregate profits specifically. Single-product firm mis-pricing raises these negative effects up to 93.3% of forecast losses, a significant jump suggesting compounding effects, not followed however by aggregate profits, which result in similar losses - 81.5% - to the optimal. In both cases, however, up to 45% of all goods are under-supplied when mis-pricing relative to the forecasted at those prices. In general, greater numbers of products and considered bundles tend to diminish the gaps between forecasts and optimal outcomes, whereas linear dependence widens them.

To demonstrate the flexibility and parsimony of linear demand over large choice sets, I then proceed to estimate a parsimonious structural model under the same specification, addressing four common concerns regarding linear demand estimation: a *curse of dimensionality* often found from quadratic growth in the number of parameters relative to the number of goods; failure to address individual choice and consumer heterogeneity distributions; and the proportionality between the number of instrumental variables necessary to address price endogeneity and the number of prices to account for (Gandhi and Nevo, 2021). To address these, I pursue an approach close to that of Pinkse and Slade (2004), estimating the matrix of cross-price elasticities semi-parametrically. To do so, I use point-of-sale data from 5 locations within a Portuguese supermarket chain from 2020 to 2023 to calculate the frequency of joint consideration – i.e. how many unique bundles were observed to include a given pair of goods

– across the product assortment. Point-of-sale data provides a repeated cross-section of the purchases of anonymised consumers, and is used to demonstrate the ability of the model to provide useful insights using less granular but more accessible data than is common in empirical Industrial Organisation.

The approach combines several advanced techniques to address challenges of nonlinearity, measurement error, and consumer heterogeneity. Frequency of joint consideration is converted into a product-by-product matrix to proxy for cross-price effects, an approach closest to that in Atalay, Frost, Sorensen, Sullivan, and Zhu (2023) and with roots in the stochastic choice literature (e.g. Manzini, Mariotti, and Ülkü, 2018). I consider two measures in the conversion process - cosine similarity a la Pellegrino (2025), and correlation a la Atalay et al. (2023) – and test each against my structural model specification. This provides a unique test in which to test different distance-based measures of product differentiation. To capture the idea that proxy values further from zero are more reliable (and those near zero are more likely to be noisy), I approximate the way in which the proxy enters the demand model via a Chebyshev polynomial series expansion. The polynomials form an orthogonal basis, reducing multicollinearity and improving the robustness and interpretability of the transformation, putting more weight on the more reliable, extreme values of the proxy. In addition, to better reflect real-world heterogeneity in consumer behavior, our strategy extends the model to include random coefficients, leading to a better overall fit and more accurate predictions when analysing demand elasticity. The estimation framework itself is implemented within a nonlinear least squares environment that is embedded in a Generalised Method of Moments (GMM) setting. Prices are instrumented using product category and supermarket fixed effects, time-varying local demographics (Fan, 2013), the average price of a given product's competitors across other supermarkets (Hausman, 1996), and the lagged number of competitors available within four increasingly narrowly defined product categories (Blundell and Bond, 1998). I form moment conditions by using these instrumental variables

and take advantage of analytical gradients to efficiently navigate the nonlinear parameter space. In doing so, I successfully recover price sensitivity parameters, which are used to measure markups across the product assortment under different ownership structures.

This paper offers significant insights for practical applications. On one hand, retailers may adjust their inventory policies by adopting models that account for joint purchase behavior. This helps balance stock levels—reducing excess inventory that ties up capital and minimises the risk of stock-outs that can lead to lost sales. On the other hand, an empirical model is introduced which can help firms deal well with consumer preference shifts over time across a large product assortment, as well as with the introduction of new goods. Pricing can be calibrated in real time with every new unique bundle observed, minimising the risk of mis-pricing.

The paper is structured as follows. Section 2 reviews the literature. Section 3 introduces the notation and definitions. Section 4 presents the theoretical framework for consumer choice over shopping bundles. Section 5 introduces the data for the demand estimation exercise and the identification strategy for the empirical specification. Section 6 concludes.

2 Literature Review

Models of demand over bundles are not new. While the Marketing literature on "shopping basket models" can be traced back to the 1990s (see Manchanda, Ansari, and Gupta (1999) for a comprehensive review), research in the context of Industrial Organisation is much more recent. From the seminal Gentzkow (2007), the topic has been picked up in by others (see e.g. Fox and Lazzati, 2017; Allen and Rehbeck, 2022; and Ershov, Laliberté, Marcoux, and Orr, 2024), with bundles invariably treated as the options in a multinomial discrete-choice model. Discrete-choice models hold substantial advantages in modelling in-

dividual choice under consumer heterogeneity and allowing for correlations in unobserved factors, while capturing realistic substitution patterns among competing products (Berry, Levinsohn, and Pakes, 1995). However, their fit for applications over large choice sets is limited due to a *curse of dimensionality*, where the number of parameters to be estimated is always more than proportional to number of goods, and therefore far too large to be manageable for a large number of products. Compounded by computational difficulties, this remains a work in progress, especially in consumer choice over bundles where these problems are particularly acute (see e.g. Iaria and Wang, 2019; and Lanier, Large, and Quah, 2022). Lastly, it is generally unclear what the inclusion of bundles adds to the intuition behind demand estimation. The present paper makes both a theoretical and an empirical contribution on this front. In modelling consumer choice over shopping bundles via quasi-linear quadratic utility and linear demand, I introduce an alternative to discrete choice models in this space, obtaining novel theoretical insights. "The unique fit of QQUMs to this particular setting is further made clear in the ease with which it accounts for large choice sets. On the empirical side, this paper offers a standard for incorporating frequency of joint purchases to approximate cross-price effects. and thus facilitate demand estimation.

Linear demand has long been seen with skepticism (Choné and Linnemer, 2020). Gandhi and Nevo (2021) provides a detailed reasoning behind the limited use linear demand has seen in empirical Industrial Organisation. In their baseline conceptualisation of the problem of demand estimation for a number of differentiated goods in a given market, they express quantity demanded of a given product as a function of own and competitors' prices, observable demand shifters, and unobserved demand shocks. The authors' substantive criticism (echoed in Berry et al. 2014) comes down to four points: the estimation of own- and cross-price effects, which the authors see as causing a curse of dimensionality; that the model fails to address individual choice and consumer heterogeneity distributions; that it cannot predict demand for new goods, even if accounting for product characteristics; and that, similar to

the quadratic increase in the number of parameters to estimate relative to the number of products, so too does the number of instrumental variables necessary to address price endogeneity increase for a greater number of prices.

This paper contributes to the applied literature which attempts to tackle these issues. Pinkse and Slade (2004) is an important contribution to this research - it tackles dimensionality concerns head on, estimating the matrix of cross-price elasticities semiparametrically, with these said to be a function of a given distance-based measure of product differentiation. While the distance measure is determined by the researcher, equal to one if a product pair is of the same type - type may be defined by product category or over a certain characteristic given a particular product space - and zero otherwise, the function is approximated via a series expansion. To instrument prices for every expansion term, both Hausman, Leonard, and Zona (1994)- and Berry, Levinsohn, and Pakes (1995)-type instruments are pursued, the number of which no longer scales with the number of goods. Pinkse and Slade (2004), however, start from a demand function based on a "normalised-quadratic indirect utility function" which is never specified by a well-defined utility maximisation problem (Gandhi and Nevo, 2022)[1]. One contribution of the present paper is, then, precisely to highlight the qualities of the QQUM to fill in that gap, in its ability to micro-found linear demand (Amir, Erickson, and Jin, 2017; for a larger discussion on this, see the Appendix). In this same vein, Dubois, Griffith, and Nevo (2014) measures product substitution via a distance-based measure on the products' characteristics directly. Pellegrino (2025) uses, as a proxy for cross-price effects, the product cosine similarities derived from textual cues identified via natural language processing (NLP) of 10-K filings of the universe of public corporations in the United States. Magnolfi, McClure, and Sorensen (2025) obtains, via an online survey, product comparisons of the kind "product A is closer to B than it is to C", which are then run through an algorithm to compute a distance-based measure of product similarity. For these

[1]According to Berry et al. (2014), this is a problem which the literature reviewed by Manchanda et al. (1999) similarly faces.

papers, the main issue is not the curse of dimensionality, but rather the data required to produce the cross-price effects matrix - Dubois et al. (2014) relies on characteristics data, which requires the researcher to determine both which characteristics matter and how the measure them; Pellegrino (2025)'s approach only works for publicly-listed, (near) single-product firms; and Magnolfi et al. (2025) presents methodological difficulties in survey scaling. Another contribution of this paper is therefore the usage of shopping bundles, rather than product characteristics, to produce the proxy for cross-price effects.

This is not to say this paper constitutes the first time that shopping bundle data is used to approach cross-price effects[2]. However, despite the QQUM specification being an obvious beneficiary of cross-price effect proxies, I have found no structural paper that has combined the two together or provided an economic test to assess their performance.

3 Model Preliminaries

In quasi-linear quadratic utility (QQUM) models, a representative consumer[3]'s utility maximisation problem is given by:

$$max_{\boldsymbol{q}}\ U(\boldsymbol{q}) = \boldsymbol{q}'\boldsymbol{\delta} - \frac{1}{2}\boldsymbol{q}'M\boldsymbol{q} + q_0 \quad \text{s.t.} \quad \boldsymbol{q}'\boldsymbol{p} + q_0 \leq M \tag{1}$$

$\mathcal{K}$ is the set of all differentiated goods $k = 1, \ldots, K$ said consumer optimises its consumption over, with $\boldsymbol{q}$ the $K \times 1$ vector of consumption per good. $\boldsymbol{\delta}$ is a strictly positive $K \times 1$ vector of marginal product utilities. The representative consumer draws utility from each product both directly and via interaction effects across those it purchases. These interaction effects between products are defined by a $K \times K$ sparse matrix M. The off-diagonal entries of the matrix may have any sign. q_0 is the numeraire[4].

[2]For recent developments in Machine Learning and Statistics see Ruiz, Athey, and Blei (2020), Tian, Lautz, Wallis, and Lambiotte (2021), and Atalay, Frost, Sorensen, Sullivan, and Zhu (2023).

[3]For a discussion of the representative consumer assumption, see the Appendix.

[4]The model satisfies local non-satiation and free disposal. Local non-satiation requires that, for any

The properties of the QQUM are well-understood[5]. M must be positive definite, as to satisfy strict concavity, downward-sloping demand, and dominant own-price effects. The model assumes no income effects[6]. The representative consumer objective function can be re-written as follows:

$$\mathcal{L}(\boldsymbol{q}) = \boldsymbol{q}'(\boldsymbol{\delta} + \phi\boldsymbol{p}) - \frac{1}{2}\boldsymbol{q}'M\boldsymbol{q} + q_0 \tag{2}$$

The (negative) term ϕ maps price sensitivity/marginal utility of income. Initial marginal utilities are assumed to be positive, i.e., $\boldsymbol{\delta} + \phi\boldsymbol{p} > 0$.

One of the most appealing features of QQUM is its ability to generate linear demand systems that remain tractable even when a market consists of a very large number of differentiated goods. Quasi-linearity allows the derived demands to depend solely on the vector of prices (without assumptions about market size or income effects) and the interactions among goods, as captured by an interaction matrix. In effect, the QQUM yields linear demand functions of the following form:

$$\boldsymbol{q} = M^{-1}(\delta + \phi\boldsymbol{p}) \tag{3}$$

with price elasticity proportional to the inverse of the interactions matrix:

$$\frac{\partial log(q_a)}{\partial log(p_b)} = \phi\{M^{-1}\}_{ab},\ \forall a, b \in \mathcal{K} \tag{4}$$

combination of goods consumption, there is always another combination arbitrarily close by that provides a strictly higher utility. The linear term in q_0 guarantees that for any given combination one can always obtain a higher utility by increasing q_0 by an arbitrarily small positive amount. q_0 also enables free disposal as it acts as a flexible numeraire that absorbs any surplus of other goods, ensuring the consumption set remains downward closed.

[5]See Amir, Erickson, and Jin (2017) and Choné and Linnemer (2020) for an exhaustive treatment.

[6]This may be restrictive in some scenarios. The reader is directed to Tirole (Introduction, p.11, 1988) and Vives (1987) for the relevant conditions.

This linear structure can be particularly useful when desirable to solve for closed-form equilibrium outcomes (e.g. due to computational efficiency gains), and scales well as the number of goods becomes large. The inversion of the matrix M is the only nontrivial computational step, and standard linear algebra techniques and software can efficiently handle high-dimensional product spaces and sparse matrices. When M is known, the model automatically "normalises" the aggregate demand without artificial boundaries between goods or via imposing, extraneous market size assumptions. This is in contrast to alternative specifications where the aggregation over many goods may require additional restrictive normalisations which may lack microeconomic justification. The M matrix can also accommodate rich patterns of substitution and complementarity among a large number of products.

4 A Theory of Consumer Choice with Bundle Choice Sets

4.1 Setup

A less well-known property of the QQUM is its ability to handle deviations from standard consumer choice. An example of this - which additionally includes issues such as a large product assortment and difficulties with defining market boundaries - is the supermarket industry. In supermarkets, consumers are understood to make consumption choices based on only a few latent factors, which drive joint purchases of goods. We may call these joint purchases *bundles*, in the spirit of Gentzkow (2007). There are many latent factors which may drive bundle purchases, such as: needs or preferences, in that some bundles may tend to include products that "go together" because they serve a common need, say, a breakfast bundle versus a dinner bundle; seasonality or promotions, where a bundle might reflect seasonal trends (holiday specials, summer outdoor items) or be influenced by a particular promotion; or category or style, where purchases might cluster by category (electronics, clothing, groceries) or even by style (premium vs. budget items).

Assume then that representative consumer restricts its purchase decisions to bundles, and that these are themselves restricted to a choice set . A choice set is the subset of a full set of available alternatives that a consumer actually pays attention to and evaluates when making a decision. Although a consumer might face a very large number of options in theory, in practice they may form a much smaller, subjective set — often due to limited information, attention, or search costs. Let $\mathcal{J}$ be the discrete set of bundles $j = 1, \ldots, J$ considered by a representative consumer, and A the $K \times J$ sparse matrix of product presence per unique bundle considered, each element a_{kj} constituting the number of units of a given good k in a given bundle j considered. $\boldsymbol{z}$ is the $J \times 1$ vector of the consumer's consumption per bundle (divisible for simplicity). We may express $\boldsymbol{q}$ as a function of $\boldsymbol{z}$:

$$\boldsymbol{q} = A\boldsymbol{z} \tag{5}$$

This identity is evocative of Lancasterian product characteristics models, with the particularity of $\boldsymbol{q}$ being on the LHS and the choice variable on the RHS. These differences evolve into an entirely different set of mathematical calculations as a result, and have different implications.

Vectors $\boldsymbol{q}$ and $\boldsymbol{z}$ have different dimensions. If a consumer makes purchase decisions over bundles of goods, the set of all possible goods consumption vectors will span the column space of matrix A, codifying the set of bundles considered by said consumer. Both the rows and columns of matrix A are unique and non-zero, but are generally not presumed linearly independent. In fact, if matrix A has more rows than columns, that is to say, if the representative consumer considers more products across bundles than the total number of bundles it considers, then linear dependency is assured (Hurewicz and Wallman, 1941). If instead there are more columns than rows, it can be shown that, if the purchase patterns

are driven by a limited number r of latent factors (the "common needs" that cause certain products to be bought together) and if r is smaller than the number of products, then the rows of A will lie in an r–dimensional subspace, and we will once again have linear dependency. Suppose there are four products—milk, bacon, eggs, and pasta—and two latent factors corresponding to the purposes "breakfast" and "dinner." Assume that for breakfast only milk and eggs are typically purchased, whereas for dinner pasta is essential, and bacon may be purchased in either context with similar frequency. We can map these preferences with a latent loading matrix C, where the rows correspond to the products (ordered top to bottom as milk, bacon, eggs, and pasta), and the columns correspond to the latent factors (breakfast and dinner, left to right):

$$C = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 2 & 0 \\ 0 & 1 \end{pmatrix} \tag{6}$$

Each matrix element, the related product's score on the associated latent factor, quantifies how strongly that product is characterised by the latent need. Next, assume that this individual's shopping trips (bundles) also reflect these purposes. For instance, suppose there are six bundles characterised by the following latent factor scores (with rows for breakfast and dinner, and columns for bundles):

$$B = \begin{pmatrix} 3 & 1 & 2 & 4 & 2 & 3 \\ 0 & 2 & 1 & 1 & 2 & 0 \end{pmatrix} \tag{7}$$

From this particular combination of weights, we can define the product-bundle matrix A:

$$A = C \cdot B = \begin{pmatrix} 3 & 1 & 2 & 4 & 2 & 3 \\ 3 & 3 & 3 & 5 & 4 & 3 \\ 6 & 2 & 4 & 8 & 4 & 6 \\ 0 & 2 & 1 & 1 & 2 & 0 \end{pmatrix}, \ rank(A) = 2 \tag{8}$$

In this example, the structure of C forces all purchase profiles (the rows of A) to be linear combinations of only two latent factors. For example, observe that the third row is

exactly twice the first row, and the fourth row can be expressed as a linear combination of the first two rows. Even though there are four products, their purchase patterns are fully explained by only two underlying needs. Thus, when bundle purchase patterns are driven by a few latent factors, the effective dimensionality of the A matrix is reduced. If the number of latent factors r is less than the number of products, the rows of A will lie (exactly or approximately) in an r–dimensional subspace. This will thereby create nontrivial linear dependencies even where there are more bundles considered than products within them[7].

Proposition 1*: If $rank(A) < K$, the vector of product consumption $\boldsymbol{q} = A\boldsymbol{z}$ is confined to the column space of A.*

In other words, $\boldsymbol{q}$ may not take arbitrary values in $\mathbb{R}$; utility maximisation over bundles limits the range of product quantities demanded which we may observe. As shown below, this may have significant implications for equilibrium outcomes. Let us now return to (2).

Through the identity (5), we can re-write the representative consumer objective function as:

$$\mathcal{L}(\boldsymbol{z}) = \boldsymbol{z}'A'(\boldsymbol{\delta} + \phi\boldsymbol{p}) - \frac{1}{2}\boldsymbol{z}'A'MA\boldsymbol{z} + z_0 \tag{9}$$

Optimising in $\boldsymbol{z}$ we obtain the following FOC:

$$(A'MA)\boldsymbol{z} = A'(\boldsymbol{\delta} + \phi\boldsymbol{p}) \tag{10}$$

[7]The likelihood of observing linear dependency will also depend on the level of aggregation the data is submitted to. At an individual level, if a single person's bundle choices are strongly influenced by just a few factors, then the purchase patterns will exhibit clear low–dimensional structure, i.e. a high degree of linear dependence. However, if data is aggregated across many individuals, each individual might be influenced by a similar set of factors but with different weights. This heterogeneity may blur the low–rank structure: the aggregated A matrix can have contributions from many individuals with differing factor loadings, which increases the effective diversity (or rank) of the aggregated data. In other words, while each individual's bundle might be low rank, averaging over many individuals could add enough independent variation that the aggregate no longer possesses said dependencies.

Expressing (10) in terms of $\boldsymbol{z}$ introduces the first problem derived from our dimensionality constraint. Unlike M,

***Lemma 1**: $A'MA$ is positive semi-definite, and therefore not generally invertible.*[8]

To proceed, I multiply both sides of the FOC by $(A'MA)^+$, for $(\cdot)^+$ the Moore-Penrose pseudoinverse (Moore, 1920; Penrose, 1955). However, on the LHS, $(A'MA)^+(A'MA) \neq I$ does not hold in general. A general form of (10) relative to $\boldsymbol{z}$ must account for the so-called *homogenous part* of differential equations:

$$\boldsymbol{z} = (A'MA)^+A'(\boldsymbol{\delta} + \phi\boldsymbol{p}) + (I - (A'MA)^+(A'MA))\boldsymbol{y} \tag{11}$$

for $\boldsymbol{y} \in \mathbb{R}$ a random vector; there are thus a multiplicity of mappings of $\boldsymbol{p}$, as the strict concavity condition we had imposed on M no longer holds for $A'MA$.

For profit-maximising firms, however, the dimensionality constraint comes to our aid. Following a re-statement of $\boldsymbol{z}$ as $\boldsymbol{q} = A\boldsymbol{z}$, the following holds.

***Proposition 2**: While the mapping from prices to the bundle-level decision $\boldsymbol{z}$ is non-unique, the resulting product demand $\boldsymbol{q} = A\boldsymbol{z}$ is unique:*

$$\boldsymbol{q} = A(A'MA)^+A'(\boldsymbol{\delta} + \phi\boldsymbol{p}) \tag{12}$$

Proof: If $\boldsymbol{q} = A\boldsymbol{z}$, $A(I - (A'MA)^+(A'MA))\boldsymbol{y} = 0$, as $I - (A'MA)^+(A'MA)$ is in the null space of $A'MA$, which itself matches the null space of A: $\boldsymbol{x}'A'MA\boldsymbol{x} = (A\boldsymbol{x})'M(A\boldsymbol{x}) = 0$ if $A'MA\boldsymbol{x} = 0$. Since our M matrix is assumed positive definite, the only way for us to have $(A\boldsymbol{x})'M(A\boldsymbol{x}) = 0$ is if $A\boldsymbol{x} = 0$. □

[8] Proofs of all Lemmas may be found in the Appendix.

The identity breaks the strict concavity assumption generally required for the well-behavedness of the utility function. Despite this, its ensuing use - following the pseudoinverse and resulting non-unique $\boldsymbol{z}$ - ultimately isolates the single unique mapping from $\boldsymbol{p}$ to $\boldsymbol{z}$ of use for obtaining $\boldsymbol{q}$.

The sparsity of the A and M matrices minimises individual product demand's reliance on the price vector of the entire product assortment (and vice-versa). Calculating the cross-price elasticities of demand, the remaining effects are held to be proportional to the expression $A(A'MA)^{-1}A'$:

$$\frac{\partial log(q_a)}{\partial log(p_b)} = \phi \frac{p_b}{q_a} \boldsymbol{a_a}'(A'MA)^{+}\boldsymbol{a_b} \lesseqgtr 0, \ \forall a, b \in \mathcal{K} \tag{13}$$

Diagonal dominance is a convenient, albeit stringent, mathematical property often used to simplify theoretical arguments such as proving uniqueness of equilibrium or establishing helpful properties for matrix inverses. $A(A'MA)^{+}A'$ will generally not be diagonal dominant even where M is. In general:

Lemma 2: *$A(A'MA)^{+}A'$ is symmetric and positive semi-definite.*

Under the dimensionality constraint, the demand system (12) still satisfies the Law of Demand [9] and features equilibrium-dependent heterogeneous elasticities of substitution. As such, it enables heterogenous markups. This is a result of those firms with higher output q_i facing less elastic (residual) demand (Pellegrino, 2025). Regardless of whether goods are complements or substitutes, the total cross-price effect will be determined both by direct interaction effects between products, but also how the remaining products across the assort-

[9] Under the usual smoothness assumptions, if the Jacobian of the demand function is negative semi–definite then the function is locally monotone non–increasing with respect to prices—that is, it satisfies the law of demand. For a positive semi-definite $A'(A'MA)^{+}A'$, $\nabla \boldsymbol{q} = \phi A(A'MA)^{+}A'$, which is negative semi-definite for $\phi < 0$.

ment interact with these in bundles.

4.2 Equilibrium

Assume first that all K differentiated products in the assortment are sold by K different single-product firms. Firms compete toward the Bertrand (Nash) equilibrium. For simplicity of exposition, marginal costs are set to $c_i = 0, \forall i \in \mathcal{K}$. The representative consumer takes prices as given.

Proposition 3*: When each of the K differentiated products is sold by a separate firm engaging in Bertrand (Nash) competition, with the representative consumer's demand given by*

$$\boldsymbol{q} = A(A'MA)^{+}A'(\boldsymbol{\delta} + \phi\boldsymbol{p})$$

and with the diagonal matrix

$$\Omega = \text{diag}(\boldsymbol{a_1}'(A'MA)^{+}\boldsymbol{a_1}, \ldots, \boldsymbol{a_K}'(A'MA)^{+}\boldsymbol{a_K}) \tag{14}$$

the unique equilibrium price vector is

$$\boldsymbol{p^*} = -\frac{1}{\phi}[\Omega + A(A'MA)^{+}A']^{-1}A(A'MA)^{+}A'\boldsymbol{\delta} \tag{15}$$

Proof: Consider a given firm i's objective function:

$$\pi_i(\boldsymbol{p}) = p_i q_i(\boldsymbol{p}) = p_i\big[\sum_j a_{ij}(A'MA)^{+}(\sum_k a_{kj}\delta_k + \phi a_{ij}p_i + \phi\sum_{k\neq i} a_{kj}p_k)\big] \tag{16}$$

The set of first-order conditions for this problem across all firms may therefore be defined

as

$$0 = A(A'MA)^{+}A'\boldsymbol{\delta} + 2\phi\Omega\boldsymbol{p} + \phi(A(A'MA)^{+}A' - \Omega)\hat{\boldsymbol{p}} \tag{17}$$

To find the Bertrand (Nash) equilibrium, we look for the fixed point $\boldsymbol{p}^*$ such that $\boldsymbol{p} = \hat{\boldsymbol{p}} = \boldsymbol{p}^*$ - i.e. that point which defines the best response as a function of all competing firms' strategies. □

Uniqueness is enabled by the following property:

Lemma 3*:* $\Omega + A(A'MA)^{+}A'$ *is invertible.*

What if all K differentiated products in the assortment are sold by a single multi-product monopolist retailer?

Proposition 4*: If a single multi-product firm sells all goods, the dimensionality constraint leads to a multiplicity of price equilibria expressed as*

$$\boldsymbol{p}^* = -\frac{1}{2\phi}\big[W\big]^{+}A(A'MA)^{+}A'\boldsymbol{\delta} + (I - W^{+}W)\boldsymbol{y} \tag{18}$$

for

$$W = \Omega + A(A'MA)^{+}A' + A(A'MA)^{+}A' \circ G \tag{19}$$

where $\boldsymbol{y} \in \mathbb{R}$ *is arbitrary and* G *is a symmetric hollow* $K \times K$ *matrix where each element* $G_{ab} = 1$ *if* $a, b \in S$, 0 *otherwise.*

Proof: A multi-product firm determines a vector of prices for its goods by maximising

the following profit function:

$$\pi_S(\boldsymbol{p}) = \sum_{i \in S} p_i [\sum_j a_{ij} (A'MA)^+ (\sum_k a_{kj}\delta_k + \phi a_{ij} p_i + \phi \sum_{k \neq i} a_{kj} p_k)] \tag{20}$$

for S the subset of products $i = 1, \ldots, K$ sold by said firm. The FOC can be expressed in matrix form as

$$0 = A(A'MA)^+ A'\boldsymbol{\delta} + \phi(2\Omega + (A(A'MA)^+ A' - \Omega))\boldsymbol{p} + \phi(G \circ A(A'MA)^+ A')\hat{\boldsymbol{p}} \tag{21}$$

To find the Bertrand (Nash) equilibrium, we operate in a manner identical to that pursued to fined the equilibrium price in the single-product firms case. □

The inclusion of an ownership matrix is problematic for the identification of a unique price equilibrium, as we cannot say anything *ex-ante* about the definiteness of G. However, for the monopoly case in particular, the multiplicity of price equilibria are payoff-equivalent. $G_{monopolist} = A(A'MA)^{-1}A' - \Omega$ for a single multi-product firm selling all goods besides the outside option, so $W_{monopolist} = 2A(A'MA)^+A'$. Given $A'(I - W^+W)\boldsymbol{y} = 0$, as discussed above, we have:

Proposition 5*: In the monopoly setting, optimal demand is given by the unique mapping*

$$\boldsymbol{q}(\boldsymbol{p}^*) = \frac{1}{2} A(A'MA)^+ A'\boldsymbol{\delta} \tag{22}$$

while the aggregate profit and consumer surplus take on the following unique close-form expressions:

$$\Pi = \sum_i^N \pi_i = \boldsymbol{q}'(\boldsymbol{p}^*)\boldsymbol{p}^* = -\frac{1}{4\phi} \boldsymbol{\delta}' A(A'MA)^+ A'\boldsymbol{\delta} \tag{23}$$

$$CS = \boldsymbol{q}'(\boldsymbol{p}^*)\boldsymbol{\delta} - \frac{1}{2}\boldsymbol{q}'(\boldsymbol{p}^*)^* M \boldsymbol{q}(\boldsymbol{p}^*) = \frac{3}{8}\boldsymbol{\delta}' A (A'MA)^+ A'\boldsymbol{\delta} \tag{24}$$

Key to the model above is the degree of linear dependency amongst the rows of A.

Lemma 4 *:If A is full row rank, then $A(A'MA)^+A' = M^{-1}$.*

In such a case, our results would match those of consumer choice over products. As this is the standard way in which the discrete-choice models common in Industrial Organisation approach consumer choice in differentiated product markets, failing to satisfy the stringent conditions in which this may apply will lead to mis-specification of equilibrium outcomes. We return to this below.

Expressions derived from this consumer choice model can be particularly useful to understand the impact on retail settings where a given products goes out of stock.

Lemma 5*: Following a product stock-out, the resulting changes in optimal demand, aggregate profits, and consumer surplus have the following closed-form expression:*

$$\Delta \boldsymbol{q}^* = -\frac{1}{4}\begin{pmatrix} \sigma_{ii}\delta_i + \boldsymbol{\sigma}_i'\boldsymbol{\delta}_{-i} \\ \boldsymbol{\sigma}_i \delta_i \end{pmatrix} \tag{25}$$

$$\Delta \Pi = \frac{1}{4\phi}\left[\sigma_{ii}\delta_i^2 + 2\delta_i(\boldsymbol{\sigma}_i'\boldsymbol{\delta}_{-i})\right] \tag{26}$$

$$\Delta CS = -\frac{3}{8}\left[\sigma_{ii}\delta^2 + 2\delta_i(\boldsymbol{\sigma}_i'\boldsymbol{\delta}_{-i})\right] \tag{27}$$

for $\Sigma = A(A'MA)^+A'$.

Even where M is symmetric and diagonally dominant (common assumptions, sufficient for positive definiteness), consumer surplus and aggregate profits may or may not *decrease* as a result of product exit, and the ways in which they do so will depend on the structure of the representative consumer's choice set. The QQUM specification aids us in allowing the expressions above to be written in closed form. This may not be the whole story however; a product whose removal raises aggregate profit may nonetheless be kept within the assortment if it can minimise the impact of other products' removal. As can be seen above, the removal of any given product is exacerbated or dampened by its interaction with others; the importance of a correct understanding of what interaction matrix to consider for appropriately addressing stock-outs is thus clear. Addressing the dimensionality constraint reveals itself fundamental for product assortment considerations.

4.3 Comparative statics

Of particular interest is the difference between model outcomes depending on the extent of dimensionality concerns raised by (5). Via regression-based comparative statics, we can quantify the relative impact of the number of linearly dependent rows, accounting for a number of factors, for the different outputs of the model.

I run a Monte Carlo simulation of product–bundle data that builds a network structure from latent factors and controls for six predictors. The dimensions of the A and M matrices were allowed to vary; each simulation generated $K \times J$ and $K \times K$ matrices with nonnegative entries, with K and J varying every simulation. Each time, the density in each matrix was also separately considered. Lastly, price sensitivity ϕ was allowed to vary. Throughout, elements of $\boldsymbol{\delta}$ were drawn from a Normal distribution with mean 10 and variance 1. A binary matrix A with r linearly-dependent rows and predetermined density is created via the latent factor approach described above. Any rows or columns that are either repeated or contain

only zeros are removed.

Under this structure I computed optimal demand, prices, consumer surplus, and aggregate profits, to explore how variations in the simulation's primitives affect these outcomes via regression analysis. Outputs were considered for both the K single-product firms and monopoly multi-product firm cases. Each time, bundle model equilibrium outcomes were compared to two alternative equilibria: (i) one elicited by a *forecast* where identity (5) is not accounted for; and (ii) a *misguided* equilibrium where prices are set to optimise over the forecast, but outcomes are considered over the correct demand responses to said mis-pricing. These are equivalent to (i) the use of M^{-1} instead of $A(A'MA)^{+}A'$ throughout, and (ii) M^{-1} for optimal pricing, followed by $A(A'MA)^{+}A'$ for the remaining computations, respectively.

The regression analysis is performed with standardised coefficients to determine (i) how the outcome variables shift in equilibrium, in terms of standard deviations, for each standard deviation increase in each predictor variable, and (ii) which predictor variables exert a stronger influence on said outcomes. The emphasis throughout is on ensuring that the simulated outcomes satisfied realistic conditions. The simulation only accepts an iteration if all computed demand vectors and prices are non-negative, have no NaN values, and the binary matrix A is non-degenerate, i.e. its smallest dimension exceeds its rank.

In the K single-product firms case, the results suggest the average optimal supply is 67% greater when forecasted without accounting for dimensionality issues. Mis-pricing as a result of such incorrect forecasts leads to similar oversupply, 59.4% higher on average than the actual optimal supply level. On welfare, consumer surplus is forecast as 32.7% higher than is found to be the case at the optimal supply level. Mis-pricing does deliver higher consumer surplus however, roughly 26.9% higher than at the profit-maximising optimum. Aggregate profits are forecast as 22.4% higher than at the optimum. Most surprisingly, perhaps is that

Table 1: Standardized Regression Results for 18 Forecast Models (7 Predictors). Robust standard errors are in parentheses. Significance levels: *p<0.10, **p<0.05, ***p<0.01. Model names are given in the notes.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
μ (%)	73.7	93.3	83.9	81.6	81.5	83.7	31.8	45.8	45.1	67.0	59.4	6.35	32.7	26.9	5.0	22.4	16.4	5.0
Intercept	5.84×10^{-17}	1.29×10^{-16}	-5.57×10^{-17}	-1.37×10^{-16}	-2.05×10^{-16}	3.44×10^{-17}	4.72×10^{-16}	-9.81×10^{-16}	3.66×10^{-16}	5.94×10^{-20}	-2.15×10^{-17}	-1.9952×10^{-17}	2.1065×10^{-16}	-4.668×10^{-17}	-1.8902×10^{-16}	4.3772×10^{-17}	4.5428×10^{-18}	-1.7802×10^{-16}
	(0.08921)	(0.09764)	(0.09687)	(0.09601)	(0.09645)	(0.09664)	(0.07306)	(0.10042)	(0.09994)	(0.06880)	(0.06777)	(0.08519)	(0.06900)	(0.06872)	(0.07659)	(0.07037)	(0.07068)	(0.07660)
K	0.24061**	0.11979	0.02380	0.19962*	0.02174	0.01631	0.18648**	0.10136	0.12075	−0.11906	−0.12110	−0.16121*	−0.14094*	−0.14404*	−0.09691	−0.11456	−0.11393	−0.09658
	(0.09927)	(0.10865)	(0.10780)	(0.10684)	(0.10733)	(0.10754)	(0.08130)	(0.11175)	(0.11121)	(0.07656)	(0.07541)	(0.09480)	(0.07678)	(0.07647)	(0.08523)	(0.07831)	(0.07865)	(0.08524)
J	0.11498	−0.07954	−0.07843	0.04503	−0.09585	−0.07596	0.23234**	−0.02244	0.03952	−0.2691***	−0.26687***	−0.23632**	−0.30492***	−0.30259***	−0.29471***	−0.29861***	−0.29244***	−0.29482***
	(0.10045)	(0.10994)	(0.10908)	(0.10810)	(0.10860)	(0.10881)	(0.08227)	(0.11307)	(0.11253)	(0.07747)	(0.07631)	(0.09592)	(0.07769)	(0.07738)	(0.08624)	(0.07924)	(0.07958)	(0.08625)
ϕ	0.03268	0.04569	0.13703	0.13406	0.10781	0.13224	0.03241	0.04684	0.03807	0.04776	0.04036	0.16498*	0.077203	0.079039	0.02198	0.06749	0.07022	0.02181
	(0.09420)	(0.10310)	(0.10229)	(0.10137)	(0.10184)	(0.10204)	(0.07715)	(0.10603)	(0.10552)	(0.07265)	(0.07156)	(0.08995)	(0.07286)	(0.07256)	(0.08087)	(0.07431)	(0.07463)	(0.08088)
M **density**	−0.06164	−0.09560	−0.10990	0.01497	−0.02158	−0.11359	−0.07429	−0.06499	−0.06874	0.02683	0.02890	0.035711	0.015873	0.0061917	0.10652	0.01243	−0.00853	0.10633
	(0.09468)	(0.10363)	(0.10282)	(0.10190)	(0.10237)	(0.10257)	(0.07755)	(0.10658)	(0.10607)	(0.07302)	(0.07193)	(0.09041)	(0.07324)	(0.07294)	(0.08129)	(0.07469)	(0.07502)	(0.08130)
A **density**	0.12989	0.10703	0.09428	0.10367	0.16909*	0.09369	−0.09596	−0.23242	−0.22503	0.05645	0.07871	−0.039451	0.0067635	0.022342	−0.09902	0.01875	0.04557	−0.09892
	(0.09306)	(0.10185)	(0.09306)	(0.10015)	(0.10061)	(0.10081)	(0.07621)	(0.10475)	(0.10425)	(0.07177)	(0.07069)	(0.08886)	(0.07198)	(0.07168)	(0.07990)	(0.07341)	(0.07373)	(0.07990)
Linearly dependent rows	−0.3194***	−0.20151*	−0.25836**	−0.16160	−0.26007***	−0.26865**	−0.63406	−0.06237	−0.08994	0.7021***	0.71311***	0.45925***	0.68301***	0.68878***	0.58233***	0.68743***	0.69159***	0.58255***
	(0.09667)	(0.10581)	(0.10498)	(0.10404)	(0.10452)	(0.10472)	(0.07917)	(0.10882)	(0.10830)	(0.07456)	(0.07344)	(0.09231)	(0.07477)	(0.07447)	(0.08300)	(0.07626)	(0.07659)	(0.08301)
Removed δ	0.00947	0.07299	0.06505	−0.09656	0.03788	0.06872	0.01852	—	—	—	—	—	—	—	—	—	—	—
	(0.09310)	(0.10190)	(0.10109)	(0.10019)	(0.10065)	(0.10085)	(0.07625)	—	—	—	—	—	—	—	—	—	—	—
Summary Statistics:																		
N										100								
RMSE	0.892	0.976	0.969	0.960	0.964	0.966	0.731	1.000	0.999	0.688	0.678	0.852	0.690	0.687	0.766	0.704	0.707	0.766
R^2	0.26	0.114	0.128	0.143	0.136	0.132	0.504	0.0629	0.0719	0.56	0.573	0.326	0.558	0.561	0.455	0.54	0.536	0.455
F (p)	4.63 (0.000179)	1.69 (0.121)	1.93 (0.0739)	2.20 (0.0411)	2.06 (0.0557)	2.00 (0.0632)	13.4 (8.9e-12)	0.882 (0.524)	1.02 (0.424)	16.7 (4.56e-14)	17.7 (1.2e-14)	6.35 (4.31e-06)	16.6 (5.89e-14)	16.8 (4.1e-14)	11 (5.35e-10)	15.4 (3.33e-13)	15.2 (4.88e-13)	11 (5.38e-10)

Notes: Model (1): `Consumer Surplus following Product Removal (Optimal v. Forecast, Firms)`; (2): `Consumer Surplus following Product Removal (Mis-pricing v. Forecast, Firms)`; (3): `Consumer Surplus following Product Removal (Optimal v. Forecast, Monopoly)`; (4): `Aggregate Profits following Product Removal (Optimal v. Forecast, Firms)`; (5): `Aggregate Profits following Product Removal (Mis-pricing v. Forecast, Firms)`; (6): `Aggregate Profits following Product Removal (Optimal v. Forecast, Monopoly)`; (7): `Degree of undersupply (%, Forecast v. Optimal, Firms)`; (8): `Degree of undersupply (%, Forecast v. Mis-pricing, Firms)`; (9): `Degree of undersupply (%, Forecast v. Optimal, Monopoly)`; (10): `Average demand (Forecast v. Optimal, Firms)`; (11): `Average demand (Mis-pricing v. Optimal, Firms)`; (12): `Average demand (Forecast v. Optimal, Monopoly)`; (13): `Consumer Surplus (Forecast v. Optimal, Firms)`; (14): `Consumer Surplus (Mis-pricing v. Optimal, Firms)`; (15): `Consumer Surplus (Forecast v. Optimal, Monopoly)`; (16): `Aggregate Profits (Forecast v. Optimal, Firms)`; (17): `Aggregate Profits (Mis-pricing v. Optimal, Firms)`; (18): `Aggregate Profits (Forecast v. Optimal, Monopoly)`.

mis-pricing by individual firms nonetheless leads to a 16.4% gain in aggregate profit relative to the optimal price equilibrium.

In the multi-product retailer/supermarket case, effects are more subdued. The average optimal supply is only 6.35% greater when forecasted without accounting for dimensionality issues. No mis-pricing occurs as a result of such incorrect forecasts, as the unique price equilibrium in that case belongs to the subset of price equilibria found when accounting for dimensionality issues. On welfare, consumer surplus is forecast as 5% higher than is found to be the case at the optimal supply level. Mis-pricing once again delivers the same surplus level as the optimal supply case. Aggregate profits are forecast at 5% greater than at the optimum, though none is lost as a result of mis-pricing.

Whether mis-pricing in fact raises aggregate profits depends on the simulated market structure, but what drives this when it does occur? Notably, production is highest in the correctly-specified monopolist multi-product retailer, followed by that amongst the single-product firms in the incorrect specification, and only then that amongst those in the correct specification. Price magnitudes go the opposite way. The finding that mis-pricing leads on average to oversupply and under-pricing is stable across all simulations, including in cases where we restrict products to being either independent or imperfect substitutes. This pattern provides important insight into how bundling influences outcomes differently from product complementarity. When complementary goods are offered together by a single firm, prices tend to be lower and sales volumes higher. In contrast, when the purchase of goods is linearly dependent, firms are inclined to set higher prices while reducing output. In other words, consumer choice over bundles appears to lessen competitive pressures.

The analysis is further extended to consider how linear dependency impacts predictions about the impact of product stock-outs. In the K single-product firms case, only 73.7%

of the forecasted consumer surplus loss following a stock-out effectively takes place. This number goes up to 93.3% under mis-pricing. This is a clear sign of compounding effects, with mis-pricing exacerbating the effects of product exit. Regarding aggregate profits, a greater percentage - 81.6% - of forecasted aggregate profit loss following a stock-out effectively takes place, while mis-pricing roughly guarantees similar losses - 81.5 %. In the multi-product retailer/supermarket case, and similarly to the firms case, only 83.7-83.9% of the forecasted consumer surplus and aggregate profit loss following a stock-out effectively takes place, respectively.

Mis-specification of consumer choice can thus generally be said to on average lead to (i) oversupply of goods and (ii) exacerbation of negative effects as a result of product stock-outs. Despite discounting dimensionality on average leading to oversupply, we ought not to dismiss the chances of the opposite taking place for at least some products. Indeed, this turns out to be a pretty significant percentage of all goods. In the firms case, when mis-pricing, an average of 45.8% of all goods are over-demanded relative to the forecasted at those prices, with the number leveled at 45.1% in the supermarket case. These numbers are significant, and reveal the extent of the exposure misdiagnosing dimensionality can have on stock-out risks.

How do model primitives affect these findings? Across most models, the number of linearly dependent rows in the A matrix is consistently highly significant with large standardised coefficients. Its positive effect suggests that increases in linear dependence are associated with higher forecast errors in demand, profit, and even supplier outcomes, and greater general over-production in demand. The size of the bundle choice set, J, is also significant in many regressions, usually with a negative coefficient, indicating an inverse relationship with the outcome measures. Model inputs generally appear to be less relevant for supermarkets, which also suffer less from mis-pricing, even if forecast error remains very high. Where it

comes to stock-outs, the results are similar. In general, greater numbers of products and considered bundles tend to diminish the gap between forecasts and optimal outcomes, whereas linear dependence widens it. Where it comes to stock-out risk, however, the signs flip, and M density comes to prominence as key to minimise the percentage of products which are underproduced.

5 Demand Estimation in Supermarkets

The QQUM is uniquely capable of providing theoretical foundations for the implications of consumer decision-making over shopping bundles. Retaining the example of the supermarket industry, I now demonstrate how such a model may be estimated over a large choice set. In doing so, I make use of the capabilities of linear demand while demonstrating how we may minimise potential concerns raised by previous research. I also explore the impact of linear dependence and our dimensionality constraint on real aggregated data.

5.1 Data

Empirical Industrial Organisation today relies substantially on scanner data[10]. In keeping with the twin goals of tractability and parsimony made possible by the QQUM specification, I instead make use of detailed point-of-sale data. Unlike conventional scanner data, all consumers are anonymised, so the dataset is a repeated cross-section, with no details on the customers themselves. Using point-of-sale as opposed to detailed scanner data offers distinct advantages in terms of data storage and privacy. Point-of-sale data records the final transaction summary rather than every individual barcode scan. This aggregation reduces data volume considerably, which eases storage and processing demands. Additionally, scanner data often contains highly granular, item-level information that — when combined

[10]See Dubois, Griffith, and O'Connell (2022) for a primer on scanner data usage in economics research.

with other data — can lead to inadvertent disclosure of sensitive consumer behavior and reveal individual customer patterns that could be exploited. Point-of-sale data, being more aggregated, is inherently less detailed, making it simpler to anonymise and secure. This also aligns better with current privacy regulations that emphasise minimal collection of personally identifiable information.

The point-of-sale data was obtained from *SONAE MC*, a leading supermarket chain in the Portuguese food retail market. *SONAE MC* is a subsidiary of *SONAE*, a multinational business group. The data set represents roughly 5% of all transactions conducted across 5 supermarkets between February 2020 to February 2023. It covers date and location of each transaction; the set of products purchased per transaction both at the product level and across 4 different product category layers; and units purchased, gross sales value, and discounts issued per product per transaction. All data is collected through the stores' scanners in the moment of the sale. The five supermarkets were selected prior to data analisys out of a list of 353 shops given their geographical distribution, perceived degree of monopolisation in its local market, and size heterogeneity. The selection along these lines was made to match the model's assumptions while including sufficient heterogeneity to briefly consider the internal validity of the model's outcomes under different conditions. The shops are differentiated by size and approach to the local communities.

We observe 343 321 unique bundles across the entire sample down to the number of units purchased per product. Store types have distinct bundle size and product statistics. Hypermarkets enable larger bundle sizes over a wider distribution, hold substantially more staple and unique products, and have a disproportional revenue share.

There are 88 835 total products across 49 major categories, 200 sub-categories, and 887 sub-sub-categories. Each product is observed an average of 382 times across the three years

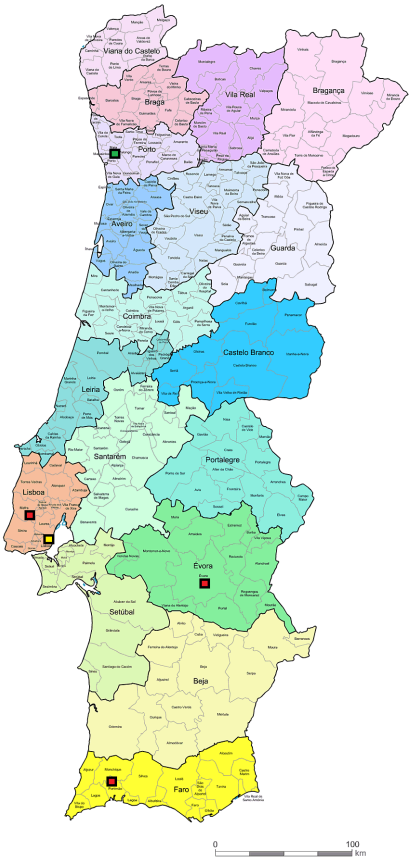


Figure 1: Shop locations - hypermarket (red), supermarket (green), concept store (yellow). Hypermarkets are major retail hubs, meant to attract consumers to a distinct location within the urban fabric. Supermarkets are mid-sized stores targeting proximity to consumers. Concept stores are experimental, cashier-less stores equipped with an array of cameras and shelf sensors meant to track a shopper's purchasing experience.

	Products	Categories
A: Bundle size (number of items)		
Mean	60.75	5.16
Standard deviation	68.94	3.78
B: Percentile bundle size		
25th	16	2
Median	36	4
75th	78	7
C: Transaction share by bundle size class		
$\leq$ 20 unique items / $\leq$ 2 groups	33.22%	30.92%
21-80 unique items / 3-8 groups	42.31%	48.80%
81-140 unique items / 9-14 groups	13.73%	16.85%
>140 unique items / >14 groups	10.73%	3.43%
D: Revenue share by bundle size class		
$\leq$ 20 unique items / $\leq$ 2 groups	8.84%	9.19%
21-80 unique items / 3-8 groups	30.38%	39.79%
81-140 unique items / 9-14 groups	22.67%	41.03%
>140 unique items / >14 groups	38.11%	9.99%

Table 2: Bundle statistics per across products and categories

Shops by Type	Mean bundle size	SD bundle size
A: Hypermarket		
Évora	71.41	79.21
Mafra	53.95	59.78
Portimão	61.00	69.26
B: Supermarket		
Maia	49.71	52.15
C: Concept		
Filipa de Vilhena	2.78	2.77

Table 3: Shop statistics

	Products	Avg Revenue	Categories	Avg Revenue
At only 1 store	43%	€89.7	4%	€13,880.8
At 2 stores	15%	€268.8	4%	€16,521.4
At 3 stores	16%	€711.8	16%	€39,737.9
At 4 stores	24%	€3,188.3	33%	€434,698.5
At all stores	2%	€8,493.6	43%	€4,381,690.7

Table 4: Product availability in stores and the average revenue brought by products at each level of availability

	#	% Unique	% Revenue	% Unique Revenue
A: Hypermarket				
Évora	49794	11.59%	31.53%	3.12%
Mafra	67603	33.79%	14.92%	6.91%
Portimão	55217	15.20%	43.57%	2.89%
B: Supermarket				
Maia	27685	3.75%	9.98%	1.62%
C: Concept store				
F Vilhena	1853	4.05%	0.02%	8.05%

Table 5: Shop statistics - note that all stores share 1671 products

of the sample, with 6, 18, and 131 observations at the 25%, median, and 75% quartiles. In other words, a substantial proportion of our sample observations come from the top 25% of all products. Sales across our sample total €99 554 423, with discounts a marginal percentage of this.

Demographic data is obtained from *PORDATA - Contemporary Portugal Database*, a free-to-use dataset developed by the Portuguese *Francisco Manuel dos Santos Foundation* which aggregates data provided by official sources, such as the *Instituto Nacional de Estatística* (*National Statistics Institute*) and *Eurostat*. *PORDATA* provides extensive detail in the following topics: Population, Health, Education, Social Protection, Employment and Labour Market, Enterprises and Personnel, Family Income and Expenditure, Housing and Comfort, Justice, Culture, National Accounts, Government Accounts, Science and Technology and Environment, Energy and Territory.

The average number of observations (i.e. purchases for consumption) per bundle is $\mu_z = 1.096$, with 5.75% of all bundles observed more than once.

The following procedures were applied to filter outliers and products for which insufficient data was made available. Any products with (1) fewer than 100 transactions, (2) sold at zero price, and/or (3) without any price variation were removed. Furthermore, for each prod-

Categories	Products	Store presence	Sales share
Groceries (Salty)	2676	5	4.64%
Groceries (Sweet)	3726	5	4.43%
Soft Drinks	1764	5	6.34%
Hygiene	2800	5	4.06%
Home Cleaning	3089	5	6.75%
Frozen goods	1035	5	2.99%
Dairy	2023	5	7.27%
Beauty	4805	5	3.42%
Essential goods	972	5	3.01%
Butchery	640	5	5.83%
Fishmonger	956	5	5.96%
Charcuterie and Cheese	1765	5	6.31%
Breakfast	2441	5	5.19%
Fruits and Vegetables	1506	5	9.30%
Bakery	1593	5	4.00%
Wine and Spirits	2330	5	4.84%
Take Away	727	5	2.09%
Luggage and Sport	1173	4	0.22%
Bio and Healthy	52	4	0.01%
Culture	14592	4	3.35%
DIY	3577	5	1.61%
Petfood and Care	1563	5	2.14%
Home Goods	10891	5	3.44%
Baby and Kids Apparel	7121	4	0.20%
Women Apparel	2890	3	0.10%
Men Apparel	2361	3	0.09%
Adult Non-Apparel	4943	4	0.21%
Nursery	558	4	0.11%
Large Domestic Products	5	3	0.01%
Small Domestic Products	197	4	0.14%
IT/Image/Sound/Photo	32	3	0.01%
TOTAL	89027		

Table 6: Product category statistics

Category	Maia	Mafra	Portimão	Lisboa	Évora
Resident population ('000,2023)	135.5	86.7	59.8	545.8	53.7
Population density (2023)	1,645.8	300.2	330.6	5,467.4	41.2
Under 15 (%, 2023)	13.9	16.6	14.7	13.4	13.5
Working age population (%, 2023)	66.8	65.6	63	63.2	63
Over 65 (%, 2023)	19.3	17.8	22.3	23.4	23.5
Foreign population (%, 2023)	2.8	7.1	20.1	19.9	4.3
Avg monthly wage (% nat. avg., 2011)	106	82	86	145	92
Unemployment (%, 2023)	5	4	11	7	5

Table 7: City statistics

uct, observations pertaining to a given month where no close competitor was found across supermarkets other than that product's are also discarded. Lastly, any products which go unobserved across a given time period in a given supermarket are assumed to not have been available for purchase, and are therefore removed. While a somewhat strong assumption, the size of our sample appear sufficiently large as to minimise potential concerns.

5.2 Empirical Strategy

From our original demand function, we estimate a nonlinear demand-model equation in which product demand $\boldsymbol{q}$ is generated by a structural function that depends on a set of observables, a nonlinear transformation of a proxy for cross-price effects, and an error term. The empirical estimation approach follows Pinkse and Slade (2004) somewhat, with some significant changes: the usage of point-of-sale bundle data instead of product types; and the use of a non-linear least-squares regression specifications to account for the implications of the QQUM functional form. I also follow Wales and Woodward (1983) in the inclusion random coefficients into a quadratic utility specification.

$$\boldsymbol{q}_t = A_t(A_t'\hat{M}_t A_t + \lambda I)^{-1}A_t' \cdot \frac{1}{S}\sum_{s=1}^{S}\left(\overline{\boldsymbol{\delta}} + \sigma_\delta \boldsymbol{\epsilon}^{(s)} + (\overline{\phi} + \sigma_\phi \eta^{(s)})\boldsymbol{p}_t\right) + \boldsymbol{u} \tag{28}$$

for $\hat{M}_t = g(W_t)$, with W a proxy for the real off diagonal interaction matrix; and $g(\cdot)$

a nonlinear function to be approximated. I allow marginal utility and price sensitivity parameters to be random, to account for consumer heterogeneity:

$$\boldsymbol{\delta}_i = \overline{\boldsymbol{\delta}} + \sigma_\delta \boldsymbol{\epsilon}_i, \quad \phi_i = \overline{\phi} + \sigma_\phi \eta_i \tag{29}$$

$\boldsymbol{\epsilon}_i$ and η_i are drawn from a standard Normal distribution. λ is a regularisation parameter - it is well-understood that

$$\lim_{\lambda \to 0} (X + \lambda I)^{-1} = (X)^+ \tag{30}$$

Despite the fact we operate in the column space of A, there may be numeric instability in practice due to the pseudoinverse. It is thus common to use some form of regularisation in practical implementations.

To proxy for M, I take advantage of the frequency with which products are considered for joint purchase - an approach closest to that in Atalay et al. (2023) and with roots in the stochastic choice literature (e.g. Manzini, Mariotti, and Ülkü, 2019). I consider two measures of product similarity - cosine similarity (see e.g. Pellegrino, 2025), and correlation (Atalay et al., 2023)[11].

One of the crucial assumptions of linear regression when performing standard estimation is that the regression error $\boldsymbol{u}$ is exogenous i.e. orthogonal to, which is to say uncorrelated with, the set of explanatory variables X: $E(X\boldsymbol{u}) = 0$ and $E(\boldsymbol{u}|X) = 0$. When this is not the case, the parameters associated with those explanatory variables are neither consistent nor unbiased, respectively. A common concern in demand estimation is that of price endogeneity due to omitted characteristics in product-level data - i.e. $E(\boldsymbol{p} \cdot \boldsymbol{u}) \neq 0$. Since it is to be expected that prices and product characteristics are choices made simultaneously by sellers during the production stage and ahead of a sale, the two may be correlated. As

[11] These measures are selected for their recent appearance in the Economics literature and straightforward implementation, though consideration of a broader number of measures may be justified.

a result, if one wishes to still use a least-squares estimator unchanged, either one assumes that the effect of those omitted characteristics is mean zero and independent across agents, or that said characteristics are orthogonal to (i.e. linearly independent from) the included characteristics. The latter assumption is rather tough to make in an environment where likely all product characteristics were chosen (close to) simultaneously but the first is rather commonplace. In such cases, sampling error tends to be unusually large, suggesting a model with low external validity and overfitting at best.

As such, this problem is usually tackled via instruments for price, which may isolate exogenous variation in the endogenous variable and therefore determine our unobserved parameters. Berry, Levinsohn, and Pakes (1995) suggest using characteristics or cost shifters of either products themselves or their competitors. Noting these are often not available, Hausman (1996) proposes using prices of the same product in geographically distinct supermarkets. In the same vein, Nevo (2001) also explores price heterogeneity, as well as market characteristics, such as geographic heterogeneity more broadly defined (e.g. local labour markets). Blundell and Bond (1998) extend this to panel data, offering the option of lagged characteristics as less impacted by potential endogeneity concerns with national pricing, centralised inventories, or chain-wide discounts and promotions. Lastly, Fan (2013) instruments on aggregate consumer or market attributes, such as education or average income, amongst others.

I take advantage of the stream of observations on product purchases across five geographically distinct supermarkets each quarter of each year in our database, as well as the repeated cross-sectional nature of our data, to define matrix Z of price instruments. I compute the average price of a given product's competitors across other supermarkets as a preliminary instrument. The (lagged) number of competitors available within 4 increasingly narrowly-defined product categories, and the (lagged) mean income across the geographical locations

of these supermarkets are additional options. These are potentially informative instruments as one may expect prices to be decreasing in the number of similarly categorised products available, and be affected by local economic dynamics. Brand fixed-effects may also be included, though for computational simplicity we use product category fixed-effects instead. Lagged average prices of competitors by product category in the same supermarket are considered. Supermarket fixed effects are similarly included as variables. The requirement that these be informative - which is to say $E(p_{it}r_{it}) \neq 0$ even after controlling for other explanatory variables present in the model is considered for all of the above. Let these be defined by a matrix Z.

The estimation procedure is carried out via a nonlinear least squares algorithm that—in our implementation—is wrapped within a Generalised Method of Moments framework. We identify the model by minimising moment conditions:

$$\min_{\boldsymbol{\theta}} ||m(\theta)||^2 = ||Z' \cdot [\boldsymbol{q} - A(A'\hat{M}A + \lambda I)^{-1}A'(\boldsymbol{\delta} + \phi\boldsymbol{p})]||^2 \tag{31}$$

where the parameter vector is $\boldsymbol{\theta} = [\boldsymbol{\delta} \quad \lambda \quad \phi \quad \boldsymbol{\gamma}]'$. Here, $\boldsymbol{\gamma}$ represents the coefficients in a Chebyshev series expansion, which we use to approximate the abovementioned $g(\cdot)$:

$$g(x) = \sum_{r=0}^{n-1} \gamma_r T_r(x) \tag{32}$$

where $T_r(x)$ denotes the r-th Chebyshev polynomial. The advantage of using Chebyshev polynomials is twofold. First, as an orthogonal basis on $[-1, 1]$ (or after appropriate rescaling), they allow for an efficient series expansion that minimises the maximum approximation error. Second, the orthogonality helps ensure that coefficients corresponding to different orders do not suffer from multicollinearity. This is particularly useful in cases where it is believed that elements of W further away from zero are more reliable proxies for the true cross-price effects (those closer to zero are more likely to be stochastic); allows the estima-

tion procedure to cleanly capture those differences without being affected by interdependent, collinear basis functions.

Closed-form expressions are identified for the Jacobian of the objective function relative to the parameters to improve computational efficiency and accelerate convergence. I employ multiple starting values to mitigate the risk of convergence to local minima and to enhance the global robustness of our estimates. After obtaining the solution, I perform diagnostic checks on the Jacobian—by verifying its rank and condition number—to ensure that our moment conditions are locally identified.

5.3 Results

Under review.

6 Conclusion

This paper develops a structural model based on quasi-linear quadratic utility and linear demand that captures how consumers make joint purchase decisions. In recognising that buyers often consider only a limited number of bundles or latent factors rather than evaluating each good independently, it offers a more realistic overview of shopping behavior in retail settings with thousands of products. By means of Monte Carlo simulations, the paper demonstrates how mis-specified consumer choice can lead to significant deviations in equilibrium outcomes; overproduction; inflated consumer surplus; and in some cases substantial aggregate profit losses. By converting frequency counts of joint product purchases into a proxy for cross-price elasticities, the paper offers a practical method for demand estimation. The reliance of this approach on accessible point-of-sale data from Portuguese supermarkets makes it a valuable tool for researchers and practitioners who might lack access to more

granular consumer data.

7 References

Allen, R., and Rehbeck, J. (2022). Latent complementarity in bundles models. Journal of Econometrics, 228(2), 322-341.

Amir, R., Erickson, P., and Jin J. (2017). On the microeconomic foundations of linear demand for differentiated products. Journal of Economic Theory, 169, 641-665.

Armstrong, M., and Vickers J. (2015). Which demand systems can be generated by discrete choice? Journal of Economic Theory, 158(A), 293-307.

Atalay, E., Frost, E., Sorensen, A., Sullivan, C., and Zhu, W. (2023). Scalable demand and markups (NBER Working Paper). National Bureau of Economic Research.

Berry, S., Khwaja, A., Kumar, V., Musalem, A., Wildur, K., Allenby, G., Anand, B., Chintagunta, P., Hanemann, W., Jeziorski, P., and Mele, A. (2014). Structural models of complementary choices. Marketing Letters, 25(3), 245-256.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. Econometrica, 63(4), 841–890.

Blundell, R., and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics, 87(1), 115–143.

Bos, I., and Vermeulen, D. (2022). On the Microfoundation of Linear Oligopoly Demand. The B.E. Journal of Theoretical Economics, 22(1), 1-15.

Chen Y., and Riordan, M. (2007). Price and variety in the spokes model. The Economic Journal, 117(522), 897-921.

Choné, P., and Linnemer, L. (2020). Linear demand systems for differentiated goods: Overview and user's guide. International Journal of Industrial Organization, 73, 1–18.

Dubois, P., Griffith, R., and Nevo, A. (2014). Do prices and attributes explain international differences in food purchases? American Economic Review, 104(3), 832–67.

Dubois, P., Griffith, R., and O'Connell, M. (2022). The use of scanner data for economics research. Annual Review of Economics, 14, 723-745.

Ershov, D., Laliberté, J., Marcoux, M., and Orr, S. (2024). Estimating Complementarity with Large Choice Sets: an Application to Mergers. Available at SSRN: https://ssrn.com/abstract=3802097

Fan, Y. (2013). Ownership consolidation and product characteristics: A study of the US daily newspaper market. American Economic Review, 103(5), 1598–1628.

Fox, T., and Lazzati, N. (2017). A note on identification of discrete choice models for bundles and binary games. Quantitative Economics, 8(3), 1021-1036.

Gandhi, A., and Nevo, A. (2021). Empirical models of demand and supply in differentiated products industries. In K. Ho, A. Hortaçsu, and A. Lizzeri (Eds.), Handbook of Industrial Organization, 4(1), 63–139. Elsevier.

Gentzkow, M. (2007). Valuing new goods in a model with complementarity: Online newspapers. American Economic Review, 97(4), 713-744.

Hausman, J. (1996). Valuation of new goods under perfect and imperfect competition. In: The Economics of New Goods, 207-248. National Bureau of Economic Research, Inc.

Hausman, J., Leonard, G., and Zona J. (1994). Competitive analysis with differentiated products. Annales d'Économie et de Statistique, 34,159-180.

Hurewicz, W., and Wallman, H. (1941). Dimension theory. Princeton University Press.

Iaria, A., and Wang, A. (2019). Identification and Estimation of Demand for Bundles. Available at SSRN: https://ssrn.com/abstract=3458543

Lanier, J., Large, J., and Quah, J. (2022). Estimating very large demand systems. Available at https://ora.ox.ac.uk/objects/uuid:13accd61-5e96-46e8-849a-fdd7c627cbfd

Magnolfi, L., McClure, J., and Sorensen, A. (2025). Triplet embeddings for demand estimation. American Economic Journal: Microeconomics, 17(1), 282–307.

Manchanda, P., Ansari, A., and Gupta, S. (1999). The "Shopping Basket": A Model for Multicategory Purchase Incidence Decisions. Marketing Science, 18(2), 95-114.

Manzini, P., Mariotti, M., and Ülkü, L. (2019). Stochastic Complementarity. The Economic Journal, 129(619) 1343-1363.

Moore, E. (1920). On the reciprocal of the general algebraic matrix. Bulletin of the American Mathematical Society, 26(9), 394–395.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. Econometrica, 69(2), 307–342.

Pellegrino, B. (2025). Product differentiation and oligopoly: A network approach. American Economic Review. Forthcoming.

Penrose, R. (1955). A generalized inverse for matrices. Proceedings of the Cambridge Philo-

sophical Society. 51(3), 406–413.

Pinkse, J., and Slade, M. (2004). Mergers, brand competition, and the price of a pint. European Economic Review, 48(3), 617-643.

PORDATA. (2025). Contemporary Portugal Database. Retrieved from http://www.pordata.pt.

Ruiz, F., Athey, S., and Blei, D. (2020). SHOPPER: A probabilistic model of consumer choice with substitutes and complements. The Annals of Applied Statistics, 14(1), 1-27.

Tian, Y., Lautz, S., Wallis, A., and Lambiotte, R. (2021). Extracting complements and substitutes from sales data: A network perspective. EPJ Data Science, 10(45).

Tirole, J. (1988). The Theory of Industrial Organization. Cambridge, Mass: The MIT Press.

Vives, X. (1987). Small Income Effects: A Marshallian Theory of Consumer Surplus and Downward Sloping Demand. Review of Economic Studies, 54(1), 87-103.

Wales, T., and Woodward, A. (1983) Estimation of consumer demand systems with binding non-negativity constraints. Journal of Econometrics, 21(3), 263-285.

8 Appendix

8.1 Proofs

Lemma 1. By construction, $\boldsymbol{x}'M\boldsymbol{x} > 0$ for all non-zero $\boldsymbol{x}$. $\boldsymbol{y}'A'MA\boldsymbol{y} \geq 0$ because, where $A\boldsymbol{y} \neq 0$, $\boldsymbol{y}'A'MA\boldsymbol{y} = \boldsymbol{x}'M\boldsymbol{x} > 0$ for non-zero $\boldsymbol{x}$. If $A\boldsymbol{y} = 0$, $\boldsymbol{x}'M\boldsymbol{x} = 0$ because $\boldsymbol{x} = A\boldsymbol{y} = 0$. □

Lemma 2. Regarding symmetry: for M positive definite and symmetric (by construction), $A'MA$ is symmetric - $(A'MA)' = A'M'A = A'MA$. By eigenvalue decomposition, $A'MA = U\Lambda U'$, and $(A'MA)^+ = U\Lambda^+U'$, where Λ^+ is formed by replacing each non-zero $\lambda, \forall i$ with $1/\lambda_i$, and leaving zeros intact. Since Λ^+ is diagonal, it is symmetric, and therefore $(A'MA)^+$ is also symmetric. Lastly, if $(A'MA)^+$ is symmetric, $A(A'MA)^+A'$ is also symmetric - $(A'(A'MA)^+A')' = A'((A'MA)^+)'A' = A(A'MA)^+A'$. Regarding positive semi-definiteness: If $A'MA$ is positive semi-definite, $(A'MA)^+$ is also positive semi-definite. As shown above, $A'MA$ is symmetric and can be diagonalised via eigenvalue decomposition, and the pseudoinverse inverts the non-zero eigenvalues while leaving the zero eigenvalues at zero. The inversion does not change the signs of the eigenvalues, so $\forall\lambda \geq 0$ still holds. If $(A'MA)^+$ is positive semi-definite, $\boldsymbol{x}'(A'MA)^+\boldsymbol{x} \geq 0$. This is the case for $\boldsymbol{x} = A'\boldsymbol{y} \neq 0$. For $\boldsymbol{x} = A'\boldsymbol{y} = 0$, $\boldsymbol{x}'(A'MA)^+\boldsymbol{x} = 0$. Therefore, $\boldsymbol{x}'(A'MA)^+\boldsymbol{x}' = \boldsymbol{y}'A(A'MA)^+A'\boldsymbol{y} \geq 0$ for all non-zero (and zero) $\boldsymbol{y}$. □

Lemma 3. The sum of a positive definite with a positive semi-definite matrix is not always invertible. By theorem, however, the sum of a positive definite Ω with a positive semi-definite matrix $A'(A'MA)^+A'$, *where the positive definite matrix is diagonal with strictly positive diagonal elements*, is itself positive definite, and therefore invertible: $\boldsymbol{x}'\Omega + A(A'MA)^+A'\boldsymbol{x} = \boldsymbol{x}'\Omega\boldsymbol{x} +$

$\boldsymbol{x}'A(A'MA)^+A'\boldsymbol{x}$ = positive+non-negative > 0. We can show that $\Omega$is strictly positive in its diagonal. Express the diagonal matrix whose diagonal matches that of $A(A'MA)^+A'$ as Ω. To guarantee strictly positive elements in the diagonal of Ω, two requirements must be satisfied. Firstly, $\boldsymbol{v}_i \neq 0, \forall i$, for $\boldsymbol{v}_i = A'\boldsymbol{e}_i$, where $\boldsymbol{e}_i$ is the standard basis vector, and $\Omega_{ii} = \boldsymbol{e}_i'A(A'MA)^+A'\boldsymbol{e}_i = \boldsymbol{v}_i'(A'MA)^+\boldsymbol{v}_i$, where $\boldsymbol{e}_i$ "picks out" the ith diagonal of Ω. This is the case as long as row i of A has at least one non-zero term, which is true of every row in A by definition, as otherwise said row would describe a product we have not observed in the data. Secondly, $\boldsymbol{v}_i'(A'MA)^+\boldsymbol{v}_i > 0, \forall i$. Now, as we mentioned earlier, $(A'MA)^+$ is positive semi-definite. However, $\boldsymbol{v}_i$ is restricted to the column space of A by construction, the same column space of $(A'MA)^+$. On the subspace where $A'MA$ is "active" (its column space), the pseudoinverse $(A'MA)^+$ behaves like a true inverse and is positive definite on that subspace. This ensures that whenever is $\boldsymbol{v}_i$ is nonzero (i.e., when it lies in the column space of A), we have $\boldsymbol{v}_i'(A'MA)^+\boldsymbol{v}_i > 0, \forall i$. Given our proof of strictly positive diagonal elements in Ω, $\Omega + A(A'MA)^+A'$ is invertible. □

Lemma 4. If A is full row rank, it can be shown that $A(A'MA)^+A' = M^{-1}$. Its singular value decomposition could be written as $A = U\Sigma V'$, for U a $K \times K$ orthogonal matrix; $\Sigma = [\Sigma_r 0]$ a $K \times J$ matrix, with Σ_r a $K \times K$ diagonal matrix containing the nonzero singular values of A, and 0 a $K \times (J-K)$ zero block; and $V = [V_1 V_2]$ is a $J \times J$ orthogonal matrix where V_1 is a $J \times K$ matrix whose columns form the orthonormal basis for the row space of A, while V_2 is a $J \times (J-K)$ matrix spanning the null space of A. A can then be re-written as $A = U\Sigma_r V_1'$. From here, $A(A'MA)^+A' = U\Sigma_r V_1'(V_1\Sigma_r U'MU\Sigma_r V_1')^+V_1\Sigma_r U' = U\Sigma_r V_1'V_1(\Sigma_r U'MU\Sigma_r)^{-1}V_1'V_1\Sigma_r U' = U\Sigma_r(\Sigma_r U'MU\Sigma_r)^{-1}\Sigma_r U' = U(U'MU)^{-1}U' = M^{-1}$ by the properties of orthogonal and diagonal matrices. □

Lemma 5. All model outputs are in one of the following forms: $\boldsymbol{y_1} = \alpha\boldsymbol{x}'\Sigma\boldsymbol{x}$ or $\boldsymbol{y_2} = \alpha\Sigma\boldsymbol{x}$ for α a strictly positive scalar, $\boldsymbol{x}$ a strictly positive vector, and Σ a positive semi-definite matrix. Suppose we wish to consider the impacts derived from the removal of the i-th component of $\boldsymbol{x}$ and Σ. These terms can be expressed as such:

$$\boldsymbol{x} = \begin{pmatrix} x_i \\ \boldsymbol{x}_{-i} \end{pmatrix}$$

and

$$\Sigma = \begin{pmatrix} \sigma_{ii} & \boldsymbol{\sigma}_i' \\ \boldsymbol{\sigma}_i & \Sigma_{-i} \end{pmatrix}$$

The notation for $\boldsymbol{\sigma}_i$ excludes the element corresponding to σ_{ii}. Starting with $\boldsymbol{y_1}$, we may restate it as:

$$\boldsymbol{y_1} = \sigma_{ii}x_i^2 + 2x_i(\boldsymbol{\sigma}_i'\boldsymbol{x}_{-i}) + \boldsymbol{x}_{-i}\Sigma{-i}\boldsymbol{x}_{-i}$$

of which the first two terms are the contribution from the i-th dimension of both $\boldsymbol{x}$ and Σ. If we remove the i-th product, i.e. the i-th dimension of $\boldsymbol{x}$ and Σ, the difference between the original and reduced term is:

$$\Delta_1 = -\alpha\left[\sigma_{ii}x_i^2 + 2x_i(\boldsymbol{\sigma}_i'\boldsymbol{x}_{-i})\right]$$

We can similarly re-write $\boldsymbol{y_2}$:

$$\boldsymbol{y_2} = \alpha\begin{pmatrix} \sigma_{ii}x_i + \boldsymbol{\sigma}_i'\boldsymbol{x}_{-i} \\ \boldsymbol{\sigma}_i x_i + \Sigma_{-i}\boldsymbol{x}_{-i} \end{pmatrix}$$

If the i-th component is dropped, the first (upper) block drops out, and the lower block's reduced

version is $\alpha\Sigma_{-i}\boldsymbol{x}_{-i}$. □

8.2 Robustness of the Representative Consumer Assumption

Quadratic utility models offer a compelling combination of analytical tractability and interpretation, yet they have long been criticised for their aggregation properties when consumers exhibit heterogeneous preferences. In the section that follows, we extend the representative consumer framework to incorporate such heterogeneity, addressing both the technical and conceptual concerns that arise when individual demand functions are aggregated. The key insight is that when each consumer's preferences are cast in quadratic form—specifically, when every consumer considers every available good—the model naturally aligns with Gorman polar form. This structure guarantees that, even with individual variations, the average demand preserves the same functional form as in the homogeneous case.

Assuming $\boldsymbol{q}_k > 0,\ \forall k \in \mathcal{K}$ for each consumer $n = 1, \ldots, N$, individual demand across goods is:

$$\boldsymbol{q}^*(\boldsymbol{p}, n) = A_n(A_n' M_n A_n)^+ A_n'(\boldsymbol{\delta} + \phi\boldsymbol{p}) \tag{33}$$

and total demand is

$$\boldsymbol{q}^*(\boldsymbol{p}) = \sum_{n=1}^{N} \boldsymbol{q}^*(\boldsymbol{p}, n) \tag{34}$$

Because QQUM is in Gorman polar form, the aggregation is exact and total demand can be restated as

$$\boldsymbol{q}^*(\boldsymbol{p}) = N \cdot A(A'MA)^+ A'(\boldsymbol{\delta} + \phi\boldsymbol{p}) \tag{35}$$

for

$$A(A'MA)^+ A' = \frac{1}{N} \sum_{n \in N} A_n(A_n' M_n A_n)^+ A_n' \tag{36}$$

We have assumed homogenous $\boldsymbol{\delta} + \phi\boldsymbol{p}$ for simplicity. For heterogenous $\boldsymbol{\delta(n)} + \phi(n)\boldsymbol{p} > 0,\ \forall n$, true by construction, this marginal utility term also averages out. The representative consumer model described throughout this paper thus approximates average bundle demand across heterogenous consumers, due to Gorman-aggregability.

Implicit in the assumption that $\boldsymbol{q}_k > 0, \forall k \in \mathcal{K}$ across all consumers is the idea that every consumer's choice set includes every available good. But suppose instead that the goods present across consumers' choice sets differ. For example, let $K - N$ be the number of goods that all consumers (the "common" goods) have in their choice set, and let N be the number of "special" goods that only a subset of consumers consider. We can still maintain a QQUM specification for each consumer by collapsing the N special goods into a single composite good S, for which

$$q_S = \sum_{j \in \mathcal{N}} \omega_j q_j \tag{37}$$

with positive weights ω_j. Now, rather than keeping separate quantities for each of the K goods, we represent the full consumer bundle as $(\boldsymbol{q}_{-S}, q_S)$. Because QQUM is already in Gorman polar form, the heterogeneity that might arise (including differences in substitution patterns among the special goods) is entirely captured in the quadratic term. In other words, all complementarity and substitution patterns appear in the M matrix. When we collapse the special goods, we must

aggregate the corresponding elements of M as follows. Suppose we partition the original M as:

$$M = \begin{pmatrix} M_{-S,-S} & M_{S,-S} \\ M_{-S,S} & M_{S,S} \end{pmatrix} \tag{38}$$

where $M_{-S,-S}$ pertains to the common goods, $M_{S,S}$ to the special goods, and $M_{S,-S} = M'_{-S,S}$ to the cross-terms. By collapsing the special goods into S, utility becomes:

$$U(\boldsymbol{q}) = \boldsymbol{q}'_{-S}\boldsymbol{\delta}_{-S} + \delta_S q_S - \frac{1}{2}\begin{pmatrix}\boldsymbol{q}_{-S} & q_S\end{pmatrix} M^* \begin{pmatrix}\boldsymbol{q}_{-S} \\ q_S\end{pmatrix} + q_0 \tag{39}$$

with δ_{-S} the weighted average of the corresponding elements in $\boldsymbol{\delta}$ for the special goods, and M^* defined as:

$$M^* = \begin{pmatrix} M_{-S,-S} & M_{S,-S}\cdot\boldsymbol{\omega} \\ (M_{S,-S}\cdot\boldsymbol{\omega})' & \boldsymbol{\omega}'\cdot M_{S,S}\cdot\boldsymbol{\omega} \end{pmatrix} \tag{40}$$

i.e. each common good's cross-effect with the composite is a weighted sum of its effects on the special goods. In other words, by "collapsing" the K special goods into one composite good through a weighted sum, we are effectively aggregating their quadratic interactions. The implication for the M matrix is that its entries related to the special goods get averaged (via the weights ω) into a single scalar (for the composite good) and into cross terms with the common goods. If every consumer is modeled this way, and if the weights (or the heterogeneity in preferences over the special goods) average out appropriately, the aggregate demand function will retain the same structural form as in the fully common case.

This may however not be convincing for the more discrete-choice-inclined IO economist. I therefore reproduce here the Choné and Linnemer (2020) generalisation of the Armstrong and Vickers (2015) result on the consistency of linear demand with a discrete-choice model where valuations are support-restricted. Let δ_{ii} be the mass of consumers interested uniquely in brand $i, i = 1, \ldots, n$, and δ_{ij} be the mass of consumers potentially interested in both good i and $j \neq i$ with $\delta_{ij} = \delta_{ji}$. Demand q has the form:

$$q_{ii} = x_{ii} - \tau_{ii}p_i \text{ and } q_{ij} = \begin{cases} x_{ij} + \tau_{ij}(p_j - p_i) & \text{if } h_{ij} < 0 \\ x_{ij} + \tau_{ij}(p_j + p_i) & \text{if } h_{ij} > 0 \end{cases} \tag{41}$$

where $\tau_{ii} > 0$ (own-price effect must be negative) and $\tau_{ij} = \tau_{ji}$ is a function of the strength of the interaction effect between the two goods, positive where $h_{ij} < 0$ and that interaction is substitution, negative when $h_{ij} > 0$ and the goods are complements. In the former case, $0 < x_{ij} < 1$ is the demand for good i when $p_i = p_j$, with $x_{ji} = 1 - x_{ij}$; in the latter, $0 < x_{ij} = x_{ji}$ are the demand for both good i and j when $p_i = p_j = 0$. In this case, we can therefore express total demand for a given product i as:

$$q_i = \sum_j \delta_{ij}q_{ij} = (\sum_j \delta_{ij}x_{ij}) - (\sum_j \delta_{ij}\tau_{ij})p_i + \sum_{j\neq i}\delta_{ij}\tau_{ij}p_j \tag{42}$$

which can alternatively be defined as $q_i = d_i - h_{ii}p_i - \sum_{j\neq i} h_{ij}p_j$ for appropriate restrictions on H. What are these and how do they limit the model? By construction, we have implicitly set $h_{ij} = -\delta_{ij}\tau_{ij}$ and $h_{ii} = \sum_j \delta_{ij}\tau_{ij}$. Combining the two, we have that $h_{ii} = \delta_{ii}\tau_{ii} - \sum_{j\neq i} h_{ij}$. Since it must be that $\delta_{ii}\tau_{ii} > 0$, this results in a requirement that $h_{ii} + \sum_{j\neq i} h_{ij} > 0$, i.e. if the prices of all goods are increased by the same amount then the demands of all goods decrease. For as

long as this condition, a weaker assumption than diagonal dominance but stronger than the Law of Demand, is satisfied, then the QQUM for a representative consumer can be microfounded by a set of heterogenous consumers distributed along a Chen and Riordan (2007)-like radial network representation of the consumer preference space with Hotelling-like utility functions for each pair of products[12].

8.3 Comparative statics under minor linear dependence

We may wish to consider what happens if we restrict our simulation to the instances where only a single row is linearly dependent. Turns out, even a small deviation in dimensionality is sufficient to spark wide deviations in competitive outcomes. In the K single-product firms case, only 89.4% of the forecasted consumer surplus loss following a stock-out effectively takes place. This number goes up to 98.6% under mis-pricing. Regarding aggregate profits, 91.5% of forecasted aggregate profit loss following a stock-out effectively takes place, but compounding effects mean 97.6% results from mis-pricing. The losses from mis-pricing instead of pursuing accurate adjustments after a stock-out are thus still significant, from roughly 6.7% to 10.3% greater losses relative to the optimal response, for aggregate profits and consumer surplus respectively. In other words, more than half of the effect on aggregate profit and more than a sixth of effects on consumer surplus come from the presence of linear dependency rather than its extent. In the multi-product retailer/supermarket case, a higher percentage - 96% - of the forecasted consumer surplus and aggregate profit loss following a stock-out effectively takes place. In almost all cases, a greater K is associated with a lower deviation from forecasts, with M making a marginal contribution to increase said deviations.

In the firms case, when mis-pricing, an average of 44.28% of all goods are over-demanded relative to the forecasted at those prices, with the number rising to 47.32% - higher still than in the general simulation - in the supermarket case. Primitives are generally uninformative.

[12]For a further discussion of this approach, see Bos and Vermeulen (2022).

Table 8: Standardised Regression Coefficients for Various Forecast Models. Robust standard errors are reported in parentheses. Significance levels: *p<0.10, **p<0.05, ***p<0.01. Full model names are given in the notes.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)
Intercept	$6.96e^{-16}$	$9.80e^{-16}$	$-5.04e^{-16}$	$2.69e^{-16}$	$5.01e^{-16}$	$2.07e^{-16}$	$-4.96e^{-16}$	$1.33e^{-16}$	$4.23e^{-16}$	$1.51e^{-16}$	$-7.97e^{-17}$	$1.04e^{-16}$	$-1.90e^{-16}$	$1.23e^{-16}$	$1.56e^{-16}$	$9.46e^{-18}$	$3.13e^{-16}$	$8.67e^{-17}$	$1.46e^{-16}$	$-4.79e^{-17}$	$9.78e^{-17}$
	(0.0921)	(0.0996)	(0.0983)	(0.0996)	(0.0983)	(0.0983)	(0.0919)	(0.1025)	(0.0924)	(0.0825)	(0.0900)	(0.0822)	(0.0885)	(0.1010)	(0.1020)	(0.0846)	(0.0843)	(0.1000)	(0.0970)	(0.0927)	(0.1000)
K	0.1045	0.1354	0.1599	0.0723	0.1261	0.1599	0.01107	−0.01710	0.2244**	−0.06623	−0.0430	−0.08567	−0.05043	0.04972	0.11558	−0.09431	−0.10266	−0.00433	−0.00259	−0.00935	−0.00433
	(0.0952)	(0.1030)	(0.1016)	(0.1030)	(0.1016)	(0.1016)	(0.0950)	(0.1060)	(0.0955)	(0.0853)	(0.0931)	(0.0850)	(0.0915)	(0.1044)	(0.1055)	(0.0875)	(0.0872)	(0.1034)	(0.1003)	(0.1003)	(0.1034)
J	0.1774*	0.04959	0.08611	0.1536	−0.01584	0.08611	0.08847	0.03765	0.2933***	−0.6029***	−0.4928***	−0.6023***	−0.5216***	−0.2059*	−0.09656	−0.5530***	−0.5549***	−0.2246**	−0.3448***	−0.4439***	−0.2246**
	(0.0977)	(0.1057)	(0.1043)	(0.1057)	(0.1043)	(0.1043)	(0.0975)	(0.1088)	(0.0980)	(0.0875)	(0.0955)	(0.0873)	(0.0939)	(0.1072)	(0.1083)	(0.0898)	(0.0894)	(0.1061)	(0.1029)	(0.1029)	(0.1061)
ϕ	0.00424	0.1404	0.1059	−0.03210	0.2013*	0.1059	−0.00824	0.02105	−0.2111*	0.09070	0.00137	0.1120	0.06148	0.04400	0.01981	0.1012	0.1178	−0.04556	0.01871	0.11965	−0.04556
	(0.0952)	(0.1030)	(0.1016)	(0.10296)	(0.1016)	(0.1016)	(0.0950)	(0.1060)	(0.0955)	(0.0853)	(0.0931)	(0.0850)	(0.0915)	(0.1044)	(0.1055)	(0.0875)	(0.0872)	(0.1034)	(0.1003)	(0.1003)	(0.1034)
M **density**	0.1003	0.05203	0.04118	0.1441	−0.01191	0.04118	−0.3874***	0.02382	0.1427	−0.04758	−0.1069	−0.05422	−0.1685*	−0.03510	−0.03380	−0.00427	−0.00427	0.00889	−0.00523	−0.02029	0.00889
	(0.0964)	(0.1043)	(0.1029)	(0.1043)	(0.1029)	(0.1029)	(0.0962)	(0.1074)	(0.0967)	(0.0864)	(0.0942)	(0.0861)	(0.0926)	(0.1057)	(0.1068)	(0.0886)	(0.0883)	(0.1047)	(0.1015)	(0.1015)	(0.1047)
A **density**	0.1196	0.1064	0.1647	0.08846	0.1649	0.1647	0.08227	−0.06701	0.05918	−0.1035	0.03700	−0.11183	0.04568	−0.06886	−0.02683	−0.1407	−0.1425	−0.03419	−0.04010	−0.04124	−0.03419
	(0.0948)	(0.1026)	(0.1012)	(0.1026)	(0.1012)	(0.1012)	(0.0947)	(0.1056)	(0.0951)	(0.0850)	(0.0927)	(0.0847)	(0.0911)	(0.1040)	(0.1051)	(0.0872)	(0.0868)	(0.1030)	(0.0999)	(0.0999)	(0.1030)
Removed δ	0.3854***	0.1126	0.1339	0.1329	0.1269	0.1339	−0.1422	0.06670	0.1220	0.04498	0.02129	0.03126	−0.01926	0.01511	0.06008	0.01447	−0.001215	0.04715	0.02747	−0.02704	0.04715
	(0.0940)	(0.1017)	(0.1003)	(0.1017)	(0.1003)	(0.1003)	(0.0938)	(0.1047)	(0.0943)	(0.0842)	(0.0919)	(0.0839)	(0.0903)	(0.1031)	(0.1042)	(0.0864)	(0.0860)	(0.1021)	(0.0990)	(0.0980)	(0.1021)
N	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
RMSE	0.921	0.996	0.983	0.996	0.982	0.983	0.919	1.03	0.924	0.825	0.900	0.822	0.885	1.01	1.02	0.846	0.843	1.00	0.970	0.927	1.00
R^2	0.204	0.0679	0.0927	0.0684	0.0932	0.0927	0.206	0.0125	0.199	0.360	0.239	0.365	0.265	0.0418	0.0220	0.327	0.333	0.0600	0.1170	0.192	0.0600
F (p)	3.96 (0.0014)	1.13 (0.351)	1.58 (0.161)	1.14 (0.346)	1.59 (0.158)	1.58 (0.161)	4.03 (0.0012)	0.197 (0.977)	3.84 (0.0018)	8.74 ($1.51e^{-7}$)	4.86 (0.00023)	8.90 ($1.13e^{-7}$)	5.59 ($5.50e^{-5}$)	0.676 (0.669)	0.348 (0.909)	7.53 ($1.35e^{-6}$)	7.72 ($9.51e^{-7}$)	0.99 (0.437)	2.05 (0.0666)	3.69 (0.0025)	0.99 (0.437)

Notes: Model (1): `Consumer Surplus following Product Removal (Optimal v. Forecast, Firms)`; (2): `Consumer Surplus following Product Removal (Mis-pricing v. Forecast, Firms)`; (3): `Consumer Surplus following Product Removal (Optimal v. Forecast, Monopoly)`; (4): `Aggregate Profits following Product Removal (Optimal v. Forecast, Firms)`; (5): `Aggregate Profits following Product Removal (Mis-pricing v. Forecast, Firms)`; (6): `Aggregate Profits following Product Removal (Optimal v. Forecast, Monopoly)`; (7): `Degree of undersupply (%, Forecast v. Optimal, Firms)`; (8): `Degree of undersupply (%, Forecast v. Mis-pricing, Firms)`; (9): `Degree of undersupply (%, Forecast v. Optimal, Monopoly)`; (10): `Average demand (Optimal v. Forecast, Firms)`; (11): `Demand variation (Optimal v. Forecast, Firms)`; (12): `Average demand (Optimal v. Mis-pricing, Firms)`; (13): `Demand variation (Optimal v. Mis-pricing, Firms)`; (14): `Average demand (Optimal v. Forecast, Monopoly)`; (15): `Demand variation (Optimal v. Forecast, Monopoly)`; (16): `Consumer Surplus (Optimal v. Forecast, Firms)`; (17): `Consumer Surplus (Mis-pricing v. Forecast, Firms)`; (18): `Consumer Surplus (Optimal v. Forecast, Monopoly)`; (19): `Aggregate Profits (Optimal v. Forecast, Firms)`; (20): `Aggregate Profits (Mis-pricing v. Forecast, Firms)`; (21): `Aggregate Profits (Optimal v. Forecast, Monopoly)`.