Non-Parametric Identification of Demand Using Variation in Product Availability*

Adam Dearing†

Cornell University and NBER

November 5, 2026

Abstract

We consider non-parametric identification of demand using micro data from multiple markets, where products' availability to consumers varies exogenously within markets. Without micro data, non-parametric identification requires instruments for both prices and quantities. Micro data can eliminate the need for quantity instruments, but existing results require variation driven by at least as many observed consumer attributes as products. We show that such variation in product availability (VPA) identifies marginal diversion ratios with respect to product quality, after which a single consumer attribute suffices. We consider VPA from three sources: second-choice data, experimental variation, and conjoint studies. Identification with second-choice data requires discrete choice and truthful reporting of second choices. Experimental variation accommodates general demand, including continuous choice and behavior other than utility maximization, but faces implementability constraints. Conjoint studies combine generality with the broad implementability of surveys, and we provide conditions under which stated preference bias does not obstruct identification.

KEYWORDS: Non-parametric identification, multi-product demand, micro data, variation in product availability, second-choice data, conjoint analysis, diversion ratios.

*We thank Daniel Ackerberg, Felipe Barbieri, Steven Berry, Jiafeng (Kevin) Chen, Christopher Conlon, Jacob Dorn, Liran Einav, Jeremy Fox, Jeff Gortmaker, Philip Haile, Lauri Kytömaa, Julie Holland Mortimer, Peter Newberry, Imke Reimers, and Kevin Williams for helpful feedback and discussions. We also thank seminar and conference participants at WashU, Yale, and the 24th Annual International Industrial Organization Conference (IIOC 2026). We thank Claude (Anthropic) for research assistance. Any errors are our own. Some results in this paper were previously circulated in early drafts of "Identifying Preference Heterogeneity in the BLP Model with Micro Data." This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

†Samuel Curtis Johnson Graduate School of Management. Email: aed237@cornell.edu.

1 Introduction

Differentiated products demand models have emerged as a workhorse in modern empirical economics.[1] However, identifying demand for differentiated products can be difficult. With market-level data alone, non-parametric identification requires instruments for both prices and quantities (Berry and Haile, 2014).[2] This places a large burden on the instruments, and weak identification is a common concern in applied work with flexible parametric models (Gandhi and Houde, 2020). Following foundational work by McFadden et al. (1977), Goldberg (1995), and Petrin (2002), a growing literature reduces the burden on instrumental variables by incorporating "micro data" that records individual consumers' observed attributes (e.g., demographics) and choices (e.g., Backus et al., 2021; Conlon and Rao, 2025; Calder-Wang and Kim, 2026). Recent advances in open-source software make it easy to include micro data in demand estimation (Grieco et al., 2025; Conlon and Gortmaker, 2026), and we expect that micro data's prevalence in applied work will continue to increase.

Berry and Haile (2024) establish non-parametric identification of demand with what we call "full-menu" micro data, where all products in the market are available to all consumers. They show how sufficiently rich variation in choices driven by observed consumer-level attributes eliminates the need to instrument for quantities. However, the required richness is substantial: there must be at least as many consumer attributes as products, and they must generate full-rank variation in quantities.[3] In many important settings, the number of products is far larger than the number of observed consumer attributes. Automobile demand is a leading example, with dozens of products but only a handful of consumer attributes (Petrin, 2002; Berry et al., 2004; Grieco et al., 2024). Other examples include ready-to-eat cereal (Backus et al., 2021), distilled spirits (Conlon and Rao, 2025), and rental housing (Calder-Wang and Kim, 2026), among many others.

[1] See, e.g., Ackerberg et al. (2007), Berry and Haile (2021), and Gandhi and Nevo (2021) for reviews of the literature. Recent applications include studies of mergers (Miller and Weinberg, 2017; Miller et al., 2021), firms' pricing conduct (Backus et al., 2021; Duarte et al., 2024; Dearing et al., 2026), tax policy (Miravete et al., 2018), and markup trends (Grieco et al., 2024; Atalay et al., 2026), among many others.

[2] Borusyak et al. (2026) show how the requirement of instruments for quantities can be softened to strong proxies for quantities when the researcher only needs to identify price counterfactuals.

[3] Berry and Haile (2024) note that this requirement is restrictive and show, in their Supplemental Appendix, that semi-parametric restrictions can reduce it to a single binary attribute. Our results keep demand non-parametric.

When consumer-attribute-driven variation is insufficient for identification, micro data containing additional information on consumers' choices can help fill in the blanks. For example, researchers can use "second-choice" data, which typically asks consumers which product they would have purchased had their chosen product been unavailable.[4] Berry et al. (2004) pioneered second-choice data's use in estimating the parametric "BLP" demand model (Berry et al., 1995), and recent applications include Grieco et al. (2024), Conlon and Gortmaker (2026), and Conlon et al. (2026). Second-choice data's counterfactual scenarios provide a type of *variation in product availability* (VPA) within markets: one product is removed while leaving the remaining products' prices, characteristics, and unobserved qualities unchanged. VPA encompasses more than second-choice data; it can involve multiple product removals, and there are several other potential sources. Stock-outs and field experiments can vary the products actually available to consumers (Conlon and Mortimer, 2010, 2013; Hickman and Mortimer, 2016), online retailers can vary the set of products each consumer sees, and a conjoint study can present respondents with the actual products in a market and randomize which set of products each respondent sees (Allenby et al., 2019).

While various types of VPA have been used to estimate parametric demand models, less is known about VPA's non-parametric identifying content. In a discrete choice setting, Conlon and Mortimer (2021) show that some types of VPA within a single market identify measures of average diversion between products. However, average diversion is not enough to non-parametrically identify demand; that requires the researcher to identify *marginal* effects, and diversion is only part of the picture. Thus, VPA's identifying content for non-parametric demand is still unclear. How far can VPA take us in non-parametrically identifying demand? And what are the tradeoffs among different types of VPA?

This paper addresses those questions. We study what VPA identifies in a general non-parametric class of demand models.[5] We show that VPA from multiple markets

[4]Second-choice data is a special case of ranked-choice data. We discuss ranked-choice data with rankings beyond the second choice in Section 4.3.

[5]Throughout, we treat demand as a primitive. Our results are robust to limited consideration models that generate demand satisfying our assumptions, but we do not separately identify preferences and consideration rules. A large literature studies that separation, using variables that shift consideration but not preferences (Goeree, 2008; Agarwal and Somaini, 2026), functional form restrictions (Abaluck and Adams-Prassl, 2021), or partial identification (Barseghyan et al., 2021a,b). See also Masatlioglu et al. (2012); Manzini and Mariotti (2014); Dardanoni et al. (2020).

point identifies *marginal* diversion ratios that capture relative changes in quantities induced by changes in unobserved product quality. With J products in the market, these diversion ratios account for all but J of the J^2 unknowns in the matrix of quality substitution effects. Once they are identified, consumer-attribute-driven variation is only needed to identify J own-quality derivatives, and a single observed attribute suffices. With the quality substitution matrix identified, there is no need to instrument for quantities. Importantly, VPA does not require consumer-attribute-driven variation to identify diversion ratios. Thus, the bulk of the substitution matrix comes from a different source of variation than in Berry and Haile (2024), who use J consumer attributes to non-parametrically identify the full substitution matrix.

We consider identification with three sources of VPA: second-choice data, experimental variation, and conjoint studies.[6] The exact conditions for identification depend on the source of VPA. Second-choice data is familiar in the multi-product demand estimation literature but faces important limitations. Its use in non-parametric identification relies on utility-maximizing discrete choice; it delivers only single-product removals; and consumers' stated second choices must coincide with their actual (counterfactual) choices when their first choice is removed from the menu.

Experimental variation, where the researcher directly manipulates product availability, uses revealed choices. It requires neither discrete choice nor utility maximization; demand may be generated by continuous choice or by behavior other than utility maximization. The offered product subsets are determined by the experiment, and the main requirements are that (i) they be assigned exogenously; and (ii) they satisfy a coverage property that we detail later. However, experimental variation is only available in settings where the products actually offered to consumers vary exogenously within markets.

Conjoint studies are a convenient alternative. Conjoints are a longstanding tool in marketing for eliciting stated preferences from menus the researcher controls (Green and Srinivasan, 1990; Allenby et al., 2019), and we consider a design that varies the availability of actual products in the market, fixing their prices and characteristics. Such conjoints share the generality of experimental variation, they give the researcher full control over the product sets offered, and they are cheap and easy to implement. The downside is that the choices in conjoints are stated rather than revealed. We

[6] While stock-outs provide VPA, they may not satisfy the conditions that our identification arguments require. We discuss this point in Section 4.2.

address this concern by providing conditions under which stated preference bias in the conjoints does not obstruct identification of actual demand. The key requirement is that respondents perceive product quality as consumers in the market do; how they map quality into stated choices can differ arbitrarily from how consumers map it into revealed choices. Intuitively, stated choices serve only to control for product quality, while revealed choices identify the diversion ratios.

We note that our results require multiple markets with VPA. In short, VPA enables us to identify markets that differ only in the unobserved quality of one product, and we identify diversion ratios by comparing quantities across such markets. So, our arguments require VPA from more than one market, regardless of its particular form.

Our results have important implications for empirical practice. First, they provide a non-parametric justification for the use of VPA in demand estimation. Thus, VPA's identifying content does not inherently rely on parametric forms, which can be viewed as tractability devices rather than essential identifying restrictions. Second, our results provide guidance on collecting VPA. Second-choice data suffices when it is available from multiple markets and the discrete choice framework applies. Otherwise, experimental VPA or conjoints over the actual products in each market provide the same identifying content, and only a handful of well-designed product subsets is needed; the required number of subsets grows slowly (logarithmically) with J.[7] Third, our results show that VPA can be used to identify demand models beyond utility-maximizing discrete choice; continuous choice and behavior other than utility maximization can be accommodated at the relatively cheap cost of running a conjoint survey in each market.

Our work is most closely related to the aforementioned Conlon and Mortimer (2021), who show that VPA from a single market identifies a weighted average of marginal diversion ratios when consumers make discrete choices. Conlon et al. (2026) build on this insight to semi-parametrically estimate a mixed logit model.[8] We differ in several respects. First, we use VPA from multiple markets to point identify the marginal diversion ratios themselves, rather than a weighted average. Second, we provide results that accommodate general demand, while they limit attention to discrete choice. Third, we focus on non-parametric identification of demand, while

[7] We discuss this combinatorial issue in more detail in Section 4.2.

[8] Conlon et al. (2026) accommodate non-parametric distributions of consumer type-specific mean utilities but still require (parametric) additive logit shocks.

Conlon et al. (2026) focus on semi-parametric mixed logit estimation. We view our results as complementary, since VPA from a single market has important practical relevance.

Our conjoint results also relate to the literature on combining revealed choice data with stated choices or beliefs.[9] Recent contributions include, *inter alia*, Feit et al. (2010), van der Klaauw (2012), Pantano and Zheng (2013), Bernheim et al. (2021), Wiswall and Zafar (2021), Giustinelli and Shapiro (2024), Meango et al. (2025), Miller et al. (2025), Briggs et al. (2026), and Norris Keiller et al. (2026). Closest to our work, Meango et al. (2025) use stated choice probabilities to control for individual-level unobserved heterogeneity when analyzing revealed choices, without any restrictions on how closely stated choices match actual (counterfactual) choices.[10] We use stated choices similarly, to control for market-level unobserved product quality. Although, the source of identifying variation is different, and our results are neither more nor less general. Their hypothetical scenarios focus on a single product (binary decision) and vary its covariates. Whereas, we vary product availability in multi-product markets, fixing covariates at their realized values. Meango et al. (2025) also consider "unmatched" data, where stated and revealed choices come from different underlying sources, and obtain bounds. Our conjoint results require that stated and revealed choices be matched in the sense that they come from the same markets.

The paper proceeds as follows. Section 2 describes the model and some preliminaries for identification, including a brief review of results from Berry and Haile (2024) for full-menu micro data. Section 3 shows how diversion ratios reduce the need for consumer-attribute-driven variation. Section 4 presents our main results on identifying diversion ratios with VPA: second-choice data in Section 4.1, experimental variation and conjoint studies in Section 4.2, and a comparison of the three sources in Section 4.3. Section 5 concludes. Throughout, we use upper case letters to denote random variables and lower case to denote realizations.

[9] Stated choices can be seen as a special case of stated beliefs, since respondents are stating what they believe they would choose if faced with the actual choice task (Manski, 1990, 1999).

[10] Pantano and Zheng (2013) use stated choice probabilities similarly in a dynamic setting, but they assume that stated probabilities of future choices match actual choice probabilities up to rounding. Outside of discrete choice, Norris Keiller et al. (2026) use firms' expectations of future inputs and outputs to control for unobserved productivity in production function estimation, allowing for some forms of biased beliefs.

2 Model and Identification Preliminaries

Much of our model setup follows Berry and Haile (2024, henceforth, "BH24"), with some important differences that we discuss in detail below. There are J products, indexed $j \in \{1, \ldots, J\}$. Markets are indexed t and are defined by realizations of prices $P_t = (P_{1t}, \ldots, P_{Jt}) \in \mathbb{R}^J$, observed product characteristics $X_t = (X_{1t}, \ldots, X_{Jt}) \in \mathbb{R}^{L \times J}$ with L characteristics per product, unobserved quality shocks $\Xi_t = (\Xi_{1t}, \ldots, \Xi_{Jt}) \in \mathbb{R}^J$, and a distribution of observable consumer-level attributes $Z_{it} \in \mathbb{R}^{\dim(Z_{it})}$, where $\dim(Z_{it}) \geq 1$ and i indexes consumers. The choice environment for consumer i in market t is then defined as $C_{it} = (Z_{it}, P_t, X_t, \Xi_t)$. The consumer chooses to purchase quantity vector $Q_{it} = (Q_{i1t}, Q_{i2t}, \ldots, Q_{iJt})$.

We denote the supports of P_t, X_t, and (P_t, X_t) as $\mathcal{P}$, $\mathcal{X}$, and $\mathcal{PX}$, respectively. The within- (and across-)market support of Z_{it} conditional on $X_t = x$ is denoted $\mathcal{Z}(x)$. The support of Ξ_t conditional on $(P_t = p, X_t = x)$ is denoted $\mathcal{E}(p, x)$. We also denote the joint supports of (Z_{it}, X_t) and (Z_{it}, P_t, X_t), respectively, as $\mathcal{ZX} = \{(z, x) : x \in \mathcal{X},\, z \in \mathcal{Z}(x)\}$ and $\mathcal{ZPX} = \{(z, p, x) : (p, x) \in \mathcal{PX},\, z \in \mathcal{Z}(x)\}$. We denote other joint supports similarly throughout the paper.

We are ultimately interested in identifying two objects: (i) the unobserved demand shocks; and (ii) the *demand function*, defined as

$$\mathcal{s}(C_{it}) = (\mathcal{s}_1(C_{it}), \ldots, \mathcal{s}_J(C_{it})) = E\left[Q_{it} \mid C_{it}\right].$$

Throughout, we refer to realizations of $\mathcal{s}(C_{it})$ as "shares." This terminology is standard in discrete choice models, and we keep it even when discussing more general demand models.

In some cases, researchers may limit attention to the *conditional demand function*,

$$\bar{\mathcal{s}}\left(Z_{it}, P_t; t\right) = \mathcal{s}\left(Z_{it}, P_t, X_t = x_t, \Xi_t = \xi_t\right),$$

which describes how demand varies with consumer attributes and prices, holding observed characteristics and unobserved shocks fixed at the realized levels in market t. Conditional demand is sufficient to identify price substitution, which is often sufficient for researchers' purposes.

Assumption 1. (Index Sufficiency) $\mathcal{s}(C_{it}) = \sigma\left(\gamma(Z_{it}, X_t, \Xi_t), P_t, X_t\right)$, where $\gamma(\cdot) =$

$(\gamma_1(\cdot), \ldots, \gamma_J(\cdot))$ and $\gamma_j(Z_{it}, X_{jt}, \Xi_{jt}) = g_j(Z_{it}, X_{jt}) + \Xi_{jt}$ for each $j \in \{1, \ldots, J\}$.

We interpret $\gamma(Z_{it}, X_t, \Xi_t)$ as a vector of product-specific qualities perceived by consumers with attributes Z_{it}. We note that prices are excluded from the index, as in BH24. Additional consumer-level attributes beyond Z_{it} can be conditioned on throughout and treated fully flexibly, including in how they interact with prices; only Z_{it} is restricted to affect demand through the quality index.[11]

For each $(p, x) \in \mathcal{PX}$, denote the support of $\gamma(Z_{it}, X_t, \Xi_t)$ conditional on $(P_t = p, X_t = x)$ as

$$\mathcal{G}(p, x) = \{g(z, x) + \xi \,:\, z \in \mathcal{Z}(x),\ \xi \in \mathcal{E}(p, x)\}.$$

Next, we impose an invertibility assumption on $\sigma(\cdot)$.

Assumption 2. (Invertibility) For each $(p, x) \in \mathcal{PX}$, $\sigma(\cdot, p, x)$ is injective on $\mathcal{G}(p, x)$.

Assumption 2 requires demand to be invertible in the quality vector; different qualities lead to different shares. This is a standard feature of workhorse demand models estimated via moment conditions, including the BLP model. E.g., Berry et al. (2013) provide sufficient conditions for invertibility that apply to many demand models. The inverse function, $\sigma^{-1}(\cdot)$, satisfies

$$\sigma^{-1}(\jmath(C_{it}), P_t, X_t) = g(Z_{it}, X_t) + \Xi_t.$$

Next, we make some support assumptions.

Assumption 3. (Support) (i) For each $x \in \mathcal{X}$: $\text{supp}\,(Z_{it}) \mid t = \mathcal{Z}(x)$ for every market t with $X_t = x$, and $\mathcal{Z}(x)$ is open and connected. (ii) For each $(p, x) \in \mathcal{PX}$: $\mathcal{E}(p, x)$ is open and connected.

Notably, Assumption 3 does *not* require large support for any of the variables. It only imposes continuous variation in both consumer attributes and unobserved market-level shocks. These are important, since our identification arguments use derivatives. Part (i) also requires that the *support* of attributes is the same in every market with the same X_t, as in BH24. The *distribution* of Z_{it} can still vary freely across markets, including in ways that depend on Ξ_t.

[11] BH24 (Supplemental Appendix S.1) relax this exclusion when Z_{it} is a vector by excluding a single coordinate from price interactions and normalizing the scale of the index on that coordinate. With a scalar Z_{it}, that coordinate is Z_{it} itself.

Assumption 3 has important implications for the support of the shares, which we now define. For $(p, x) \in \mathcal{PX}$ and $\xi \in \mathcal{E}(p, x)$, define

$$\mathcal{S}(p, x, \xi) = \{\sigma(g(z, x) + \xi, p, x) : z \in \mathcal{Z}(x)\} \qquad \text{and} \qquad \mathcal{S}(p, x) = \bigcup_{\xi \in \mathcal{E}(p,x)} \mathcal{S}(p, x, \xi).$$

$\mathcal{S}(p, x, \xi)$ is the support of share vectors attainable *within* a market where $(P_t, X_t, \Xi_t) = (p, x, \xi)$. $\mathcal{S}(p, x)$ is the combined support *within and across* all markets with $(P_t, X_t) = (p, x)$. We also define the joint support $\mathcal{SPX} = \{(s, p, x) : (p, x) \in \mathcal{PX},\ s \in \mathcal{S}(p, x)\}$.

Together, parts (i) and (ii) of Assumption 3 imply $\text{supp}\,(Z_{it}, \Xi_t) \mid (P_t = p, X_t = x) = \mathcal{Z}(x) \times \mathcal{E}(p, x)$. So, every $s \in \mathcal{S}(p, x)$ is attained by some consumer attribute value in some market with $(P_t, X_t) = (p, x)$. Assumption 3 also implies that $\mathcal{S}(p, x)$ is open and connected.[12]

Next, we impose some technical conditions.

Assumption 4. (Technical Conditions) For all $(p, x) \in \mathcal{PX}$: (i) $\sigma(\cdot, p, x)$ is continuously differentiable; (ii) $\partial\sigma(\gamma, p, x)/\partial\gamma$ is non-singular a.s. on $\mathcal{G}(p, x)$; and (iii) $\partial\sigma_j(\gamma, p, x)/\partial\gamma_j \neq 0$ on $\mathcal{G}(p, x)$ for each $j \in \{1, \ldots, J\}$.

Most of the conditions in Assumption 4 are also used in BH24. Only part (iii) of Assumption 4 is new; it ensures that diversion ratios w.r.t. quality changes (defined in Section 3) are well-defined. We note that we do not need to sign these derivatives, although they are positive in the workhorse "BLP" demand model (Berry et al., 1995).

While our setup and assumptions mostly follow those of BH24, there are some important differences beyond those already discussed. The main difference is in the number of consumer attributes and their effect on $g(\cdot)$. BH24 require $\dim(Z_{it}) = J$ and that $g(\cdot, x)$ is invertible in Z. Whereas, our use of VPA allows us to relax these assumptions; there can be as little as a single attribute ($\dim(Z_{it}) = 1$).

A second difference is that we impose a stronger index restriction than BH24, requiring the index to be a vector of product-specific quality indices. Each product's index depends only on attributes Z_{it} and its own market-level variables (X_{jt}, Ξ_{jt}).

[12] Fix $z \in \mathcal{Z}(x)$. The map $\xi \mapsto \sigma(g(z, x) + \xi, p, x)$ is continuous (Assumption 4) and injective (Assumption 2) from the open set $\mathcal{E}(p, x)$ into $\mathbb{R}^J$, so its image is open by invariance of domain. $\mathcal{S}(p, x)$ is a union of such images, so it is open as well. It is connected because it is the continuous image of the connected set $\mathcal{Z}(x) \times \mathcal{E}(p, x)$. Note that with a scalar attribute, the openness of $\mathcal{S}(p, x)$ has to come from $\mathcal{E}(p, x)$; each $\mathcal{S}(p, x, \xi)$ is then a curve in $\mathbb{R}^J$, rather than an open set.

This is a substantial assumption, although it fits well with the interpretation of the index as a product's (consumer-attribute-specific) quality. The assumption is satisfied in workhorse demand models, and we give an example in Section 2.1.

Third, we make a stronger assumption on the support of the demand shocks. BH24 only require open and connected $\mathcal{E}(p,x)$ at one price vector for each x, because their within-market variation in Z_{it} carries the argument to every other (p,x) when $\dim(Z_{it}) = J$. Whereas, that route is unavailable with a single consumer attribute, so we require $\mathcal{E}(p,x)$ to be open and connected at *every* (p,x) (Assumption 3(ii)). This is a cost of reducing $\dim(Z_{it})$: relative to BH24, we ask less of the consumer attributes and more of the cross-market variation in Ξ_t.

Assumptions 1-4 accommodate both discrete and continuous-choice demand. Next, we describe the discrete choice setting in more detail. The discrete choice framework underlies the design of second-choice surveys, and we rely on this framework to interpret second-choice data in Section 4.1.

2.1 Discrete Choice

Consumers make a discrete choice among J products, indexed $j \in \{1,\dots,J\}$, and an outside option indexed $j = 0$. Each consumer has a vector of utilities from consuming each product. Utility for the J products is

$$U_{it}(C_{it}) = (U_{i1t}(C_{it}),\dots,U_{iJt}(C_{it})),$$

where $U_{it}(\cdot)$ is a random function. We normalize the utility of the outside option to $U_{i0t} = 0$ without loss of generality.[13]

The consumer chooses the product with the highest utility. We assume that ties occur with probability zero, conditional on C_{it}, and proceed as if there are no ties.[14] Quantities are $Q_{ijt} = 1$ if $j = \arg\max_{j' \in \{0,1,\dots,J\}} \{U_{ij't}(C_{it})\}$ and $Q_{ijt} = 0$ otherwise.

Assumption 1′. (Index Sufficiency in Discrete Choice) Consumers make utility-maximizing discrete choices, where (i) $U_{ijt}(C_{it}) = \tilde{U}_{ijt}(\gamma_j(Z_{it}, X_{jt}, \Xi_{jt}), P_{jt}, X_{jt})$ with $\gamma_j(Z_{it}, X_{jt}, \Xi_{jt}) = g_j(Z_{it}, X_{jt}) + \Xi_{jt}$ for each $j \in \{1,\dots,J\}$; and (ii) $\tilde{U}_{it}(\cdot) \perp\!\!\!\perp C_{it}$.

[13] For an alternative utility $(V_{i0t}(C_{it}), V_{i1t}(C_{it}),\dots,V_{iJt}(C_{it}))$, we can define $U_{ijt}(C_{it}) = V_{ijt}(C_{it}) - V_{i0t}(C_{it})$ for $j = 0,1,\dots,J$, which results in $U_{i0t}(C_{it}) = 0$. Both utility functions yield the same ranking of products because $V_{ijt}(C_{it}) > V_{ij't}(C_{it})$ if and only if $U_{ijt}(C_{it}) > U_{ij't}(C_{it})$.

[14] This is a standard assumption in discrete choice demand models.

Assumption 1′ implies Assumption 1 in the discrete choice setting: demand depends on C_{it} only through $\left(\gamma_j(\cdot), P_{jt}, X_{jt}\right)_{j=1}^{J}$, since $\tilde{U}_{it}(\cdot)$ is independent of C_{it}. It imposes two restrictions on the utility function. First, utility for product j is a function of only consumer attributes and j-specific market-level variables. Second, Z_{it} and Ξ_t enter the utility function through the quality index $\gamma_j(\cdot)$ that is linear in Ξ_t. We note that Assumption 1′ also delivers index sufficiency when any subset of products is offered, since the utility function does not depend on which products are available. This property has an important role in our analysis of second-choice data in Section 4.1.

Example ("BLP" Demand): Here, we describe a generic workhorse specification of the "BLP" random coefficients demand model (Berry et al., 1995) and show that it satisfies Assumption 1′, along with our other assumptions. There are L characteristics, with $X_{jt} = \left(X_{jt}^{(1)}, \ldots, X_{jt}^{(L)}\right)$. Utility for product $j \in \{1, \ldots, J\}$ is specified as[15]

$$U_{ijt}(C_{it}) = \underbrace{h^{(0)}(Z_{it}) + \nu_{it}^{(0)}}_{\equiv \beta_{it}^{(0)}} + \sum_{\ell=1}^{L} \underbrace{\left(h^{(\ell)}(Z_{it}) + \nu_{it}^{(\ell)}\right)}_{\equiv \beta_{it}^{(\ell)}} X_{jt}^{(\ell)} - \alpha_{it} P_{jt} + \Xi_{jt} + \epsilon_{ijt} - \epsilon_{i0t}.$$

The terms $(\beta_{it}, \alpha_{it})$ are random coefficients that can be decomposed into two components: (i) consumer-attribute-driven variation in tastes for (non-price) characteristics determined by $h(\cdot)$; and (ii) random terms (ν_{it}, α_{it}) satisfying $(\nu_{it}, \alpha_{it}) \perp\!\!\!\perp C_{it}$.[16] The utility shocks $\{\epsilon_{ijt}\}_{j=0}^{J}$ are distributed i.i.d. Type I Extreme Value (Gumbel).

To see why Assumption 1′ is satisfied, we can re-write the utility function as

$$U_{ijt}(C_{it}) = \underbrace{h^{(0)}(Z_{it}) + \sum_{\ell=1}^{L} h^{(\ell)}(Z_{it}) X_{jt}^{(\ell)} + \Xi_{jt}}_{\equiv g_j(Z_{it}, X_{jt}) + \Xi_{jt}} + \nu_{it}^{(0)} + \sum_{\ell=1}^{L} \nu_{it}^{(\ell)} X_{jt}^{(\ell)} - \alpha_{it} P_{jt} + \epsilon_{ijt} - \epsilon_{i0t}.$$

With $\left(\nu_{it}, \alpha_{it}, \{\epsilon_{ijt}\}_{j=0}^{J}\right) \perp\!\!\!\perp C_{it}$, this form of utility function satisfies both conditions in Assumption 1′. Thus, the BLP model satisfies Assumption 1. It also satisfies the invertibility requirement of Assumption 2.[17] Assumption 3 is a statistical condition

[15] This specification is in line with our restriction that $U_{i0t}(C_{it}) = 0$.

[16] Consumer attributes excluded from Z_{it} can interact with price. We implicitly condition on any such attributes.

[17] Berry (1994), Berry et al. (1995), and Berry et al. (2013) provide proofs of invertibility, and

that does not depend on the demand function. The BLP model also satisfies the technical conditions in Assumption 4. Thus, the class of models satisfying Assumptions 1-4 includes BLP demand.

2.2 Identification Preliminaries

Here, we discuss some preliminaries for identification. We begin with the primary objects to be identified by micro data from many markets. We then review BH24's identification results using consumer-attribute-driven variation in full-menu choices.

In each market, the researcher observes the realization of (P_t, X_t), as well as realizations of (Q_{it}, Z_{it}) for many consumers within the market. Throughout, we adopt a standard convention for identification analysis: we treat the population joint distribution of the observables as known, and we identify markets with points in $\text{supp}\,(P_t, X_t, \Xi_t)$. So, for each (p, x, ξ) in that support, there is a market t with $(P_t, X_t, \Xi_t) = (p, x, \xi)$.

We define the consumer-attribute-specific share in market t as

$$\bar{s}_t(z) = E\left[Q_{it} \mid Z_{it} = z;\, t\right].$$

This is identified directly from market t's micro data, which contains consumer-level (Q_{it}, Z_{it}) observations for many consumers. The function $\bar{s}_t(\cdot)$ is a purely data-driven object; its identification does not rely on our assumptions.

The key demand object to be identified with micro data is the inverse-demand Jacobian w.r.t. shares,

$$\frac{\partial \sigma^{-1}(\cdot)}{\partial s} \tag{1}$$

on $\mathcal{SPX}$. By the Inverse Function Theorem, this Jacobian is equal to the multiplicative inverse of the quality substitution matrix $\partial\sigma/\partial\gamma$. Once the key Jacobian in (1) is identified, arguments from BH24 (Sections 4.2 and 4.3) can be used to identify the demand and conditional demand functions via price instruments, W_t, such that (i) W_t is excluded from the demand function; and (ii) $(P_t, X_t) \mid W_t, X_t$ is complete.[18] If both the instruments and the characteristics are exogenous, with $E\left[\Xi_t \mid W_t, X_t\right] = 0$,

Berry et al. (1995) provide a numerical inversion algorithm.

[18] $(P_t, X_t) \mid W_t, X_t$ is complete if $E\left[B(P_t, X_t) \mid W_t, X_t\right] = 0$ a.s. implies $B(P_t, X_t) = 0$ a.s. See, e.g., Newey and Powell (2003).

then both the demand and conditional demand functions are identified.[19] When only the instruments are (conditionally) exogenous, with $E[\Xi_t \mid W_t, X_t] = h(X_t)$ for some (unknown) $h(\cdot)$, the conditional demand function is still identified but the demand function is not. Importantly, there is no need to instrument or proxy for shares, which stands in contrast to settings where the researcher only has market-level data (Berry and Haile, 2014; Borusyak et al., 2026).

For the reader's convenience, we present a version of BH24's argument for identification of the full demand function. Suppose (i) Assumptions 1-4 hold; (ii) the inverse-demand Jacobian in (1) is identified; (iii) $(P_t, X_t) \mid W_t, X_t$ is complete; and (iv) $E[\Xi_t \mid W_t, X_t] = 0$. Define $\Delta\sigma^{-1}(s, p, x) = \sigma^{-1}(s, p, x) - \sigma^{-1}(s^0(p, x), p, x)$ for some $s^0(p, x) \in \mathcal{S}(p, x)$ chosen by the researcher. With $\partial\sigma^{-1}(\cdot)/\partial s$ identified and $\mathcal{S}(p, x)$ open and connected (Assumption 3), $\Delta\sigma^{-1}(\cdot)$ is identified via path integration. Fixing some $z^0(x) \in \mathcal{Z}(x)$ – also chosen by the researcher – for each $x \in \mathcal{X}$, we then obtain a market-level non-parametric instrumental variables regression of the form

$$Y_{jt} = f_j(P_t, X_t) + \Xi_{jt}$$

where

$$\begin{aligned} Y_{jt} &\equiv \Delta\sigma_j^{-1}\left(\bar{s}_t\left(z^0(X_t)\right), P_t, X_t\right) \\ f_j(P_t, X_t) &\equiv g_j\left(z^0(X_t), X_{jt}\right) - \sigma_j^{-1}\left(s^0(P_t, X_t), P_t, X_t\right). \end{aligned}$$

Here, Y_{jt} is already identified, so we can identify $f_j(\cdot)$ via non-parametric instrumental variables regression (Newey and Powell, 2003).[20] We then identify the unobserved demand shocks as $\Xi_{jt} = Y_{jt} - f_j(P_t, X_t)$ for each $j \in \{1, \ldots, J\}$. Having now identified the unobserved component of the choice environment, $C_{it} = (Z_{it}, P_t, X_t, \Xi_t)$, we can identify $\jmath(C_{it})$ directly from the micro data on consumers' choices.

Moving forward, our main focus is on identifying the inverse-demand Jacobian in (1). When only full-menu micro data is available, BH24 show that the key inverse-demand Jacobian can be identified by sufficiently rich consumer-attribute-driven variation in shares.

[19] Identification of conditional demand then follows trivially from identification of demand.

[20] The argument we use here is slightly different than BH24's. They identify $g(\cdot)$ on $\mathcal{Z}\mathcal{X}$ en route to identifying the inverse-demand Jacobian and incorporate it into Y_{jt} instead of $f_j(P_t, X_t)$. However, there is no material impact on the non-parametric IV argument, since Y_{jt} is identified from the micro data in either case.

Claim 1 (BH24, Corollary 1)**.** *Suppose Assumptions 1-4 hold. Additionally, suppose that (i)* $\dim(Z_{it}) = J$*; (ii) for all* $(x, \xi) \in supp(X_t, \Xi_t)$*,* $\gamma(\cdot, x, \xi)$ *is injective; (iii) for all* $x \in \mathcal{X}$*,* $g(\cdot, x)$ *is uniformly continuous and continuously differentiable, with* $\partial g(z, x)/\partial z$ *non-singular a.s. on* $\mathcal{Z}(x)$*; and (iv) (location and scale normalizations) for each* $x \in \mathcal{X}$*, normalize* $g(z^0(x), x) = 0$ *and* $\partial g\left(z^0(x), x\right)/\partial z = I$ *for some* $z^0(x) \in \mathcal{Z}(x)$ *at which* $\partial g\left(z^0(x), x\right)/\partial z$ *is non-singular.*[21] *Then,* $\partial\sigma^{-1}(\cdot)/\partial s$ *is identified on* $\mathcal{SPX}$*.*

We refer the reader to BH24 for a rigorous proof of the result in Claim 1. The requirement for consumer-attribute-driven variation in shares is encapsulated by non-singularity of $\partial g(\cdot)/\partial z$, which requires $\dim(Z_{it}) = J$. BH24's argument therefore needs as many attributes as products. As discussed in the introduction, there are many relevant settings where this restriction is prohibitive.

We note that Assumption 3(ii), which strengthens BH24's non-degeneracy condition (their Assumption 5), also allows an argument with a single consumer attribute and no VPA. However, its requirements still grow rapidly with J.

Claim 2. *Suppose Assumptions 1-4 hold and* Z_{it} *is scalar-valued. Additionally, suppose that (i) for each* $x \in \mathcal{X}$*,* $g(\cdot, x)$ *is* J *times continuously differentiable, with the* $J \times J$ *matrix* $\left[\partial g(z, x)/\partial z, \ldots, \partial^J g(z, x)/\partial z^J\right]$ *non-singular a.s. on* $\mathcal{Z}(x)$*; and (ii) (scale normalization) for each* $x \in \mathcal{X}$*, normalize* $\left[\partial g\left(z^0(x), x\right)/\partial z, \ldots, \partial^J g\left(z^0(x), x\right)/\partial z^J\right] = I$ *for some* $z^0(x) \in \mathcal{Z}(x)$ *at which this matrix is non-singular.*[22] *Then,* $\partial\sigma^{-1}(\cdot)/\partial s$ *is identified on* $\mathcal{SPX}$*.*

We prove Claim 2 in Appendix A.2. The idea is as follows. Fix (p, x) and $s \in \mathcal{S}(p, x)$. Under Assumption 3(ii), there is an open set of attribute values that

[21] The location and scale normalizations are without loss of generality (BH24, Section 2.5). BH24 choose $z^0(x)$ such that $\partial\bar{s}_t\left(z^0(x)\right)/\partial z$ is non-singular in some market t with $X_t = x$, which is observable and implies non-singularity of $\partial g\left(z^0(x), x\right)/\partial z$. By assumption, almost every point in $\mathcal{Z}(x)$ qualifies.

[22] This is the counterpart of the scale normalization in Claim 1(iv), with the first J derivatives w.r.t. a single consumer attribute in place of the first derivatives w.r.t. J attributes. Both are without loss of generality when any non-singular linear transformation of the index is admissible, as in BH24 (Section 2.5). Note that only product-by-product re-scalings are admissible in our setting (Appendix B), so the normalized representation may lie outside the class of Assumption 1. This does not matter for identifying demand: $\jmath(\cdot)$ is the same in every representation, and the instrumental variables argument above goes through with the transformed shocks in place of Ξ_t, since a non-singular linear transformation of Ξ_t that depends only on X_t preserves $E\left[\Xi_t \mid W_t, X_t\right] = 0$.

attain s in markets with $(P_t, X_t) = (p, x)$. Each such market reveals $d\bar{s}_t(z)/dz = M\,\partial g(z,x)/\partial z$ for a different z (the attribute value that attains s in that market), where M is equal to $\partial\sigma(\gamma, p, x)/\partial\gamma$ evaluated at $\gamma = \sigma^{-1}(s, p, x)$ and is the same in all of them. Condition (i) ensures that any open set of attribute values contains J values at which the vectors $\partial g(z,x)/\partial z$ are linearly independent. So, J such values can play the role of BH24's J consumer attributes, and they identify M up to the unknown $J \times J$ matrix formed by these vectors. The scale normalization in condition (ii) then pins down that matrix, in the same way that the normalization in Claim 1(iv) pins down $\partial g\left(z^0(x), x\right)/\partial z$ in BH24.

The condition on the derivatives in Claim 2 is the counterpart of BH24's requirement of J consumer attributes. Claim 1(iii) requires the $J \times J$ matrix of first derivatives of $g(\cdot, x)$ w.r.t. J attributes to be non-singular. Claim 2 requires the $J \times J$ matrix of the first J derivatives w.r.t. a single attribute to be non-singular. So, the required richness of attribute-driven variation still grows quickly with J.

Both conditions are strong, and they can be surprisingly hard to satisfy. The BLP model from Section 2.1 illustrates this.

Remark 1. Suppose that demand is generated by the BLP model in Section 2.1 with $J > L + 1$. Then, for any number of consumer attributes and any functions $h(\cdot)$, the rank conditions in Claim 1(iii) and Claim 2(i) fail: the $J \times J$ matrices in those conditions have rank no greater than $L + 1$.

Remark 1 holds because consumer attributes enter each product's quality index only through coefficients on a constant and the L characteristics. Every derivative of $g(z, x)$ w.r.t. z must then be a linear combination of ι and $x^{(1)}, \ldots, x^{(L)}$, where ι is a J-vector of ones and $x^{(\ell)}$ collects the ℓ'th characteristic of each product. So, the columns of the $J \times J$ matrices of derivatives in Claims 1 and 2 are spanned by $L + 1$ vectors, resulting in matrix rank no larger than $L + 1$. We note that $J > L + 1$ is typical in BLP applications; for example, automobile demand has dozens of products and only a handful of characteristics (Berry et al., 1995, 2004; Grieco et al., 2024).

Remark 1 uncovers a surprising limitation of the rank conditions required by Claims 1 and 2. If, unbeknownst to the researcher, demand were generated by a BLP model with more products than characteristics, then neither argument in Claims 1 or 2 could be used to identify demand. This limitation arises regardless of how many

consumer attributes enter $h(\cdot)$ or however non-linearly they enter.[23]

Altogether, we see that the richness of consumer-attribute-driven variation required for identification with full-menu micro data can be restrictive. VPA helps alleviate the burden on attribute-driven variation by identifying a type of diversion ratio, reducing the required attribute-driven variation from rank J to rank 1. In the next section, we define the relevant diversion ratio and elucidate its role in identification.

3 How Diversion Ratios Help

Define the *diversion ratio* from product k to j (with $k \neq j$) w.r.t. the quality index as

$$\tilde{D}_{k\to j}(\gamma, p, x) \equiv \frac{\partial \sigma_j(\gamma, p, x)/\partial \gamma_k}{\partial \sigma_k(\gamma, p, x)/\partial \gamma_k}.$$

This diversion ratio captures the relative changes in the shares for products k and j induced by a change in the quality of product k.[24]

Eventually, in Section 4, we will show how micro data with VPA identifies these diversion ratios. But first, we show how this helps identify the key inverse-demand Jacobian from (1). To build intuition, recall that $\partial \sigma^{-1}(s, p, x)/\partial s$ is the multiplicative inverse of the quality substitution matrix at quality $\gamma = \sigma^{-1}(s, p, x)$. The quality substitution matrix can be decomposed into the product of (i) a matrix with ones on the diagonals and diversion ratios on the off-diagonals; and (ii) a diagonal matrix of own-quality derivatives:

$$\underbrace{\begin{bmatrix} \frac{\partial \sigma_1}{\partial \gamma_1} & \cdots & \frac{\partial \sigma_1}{\partial \gamma_J} \\ \vdots & \ddots & \vdots \\ \frac{\partial \sigma_J}{\partial \gamma_1} & \cdots & \frac{\partial \sigma_J}{\partial \gamma_J} \end{bmatrix}}_{J\times J \text{ unknowns}} = \underbrace{\begin{bmatrix} 1 & \tilde{D}_{2\to 1} & \cdots & \tilde{D}_{J\to 1} \\ \tilde{D}_{1\to 2} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \tilde{D}_{J\to (J-1)} \\ \tilde{D}_{1\to J} & \cdots & \tilde{D}_{(J-1)\to J} & 1 \end{bmatrix}}_{J\times (J-1) \text{ unknowns}} \times \underbrace{\begin{bmatrix} \frac{\partial \sigma_1}{\partial \gamma_1} & 0 & \cdots & 0 \\ 0 & \frac{\partial \sigma_2}{\partial \gamma_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \frac{\partial \sigma_J}{\partial \gamma_J} \end{bmatrix}}_{J \text{ unknowns}}.$$

[23] The remark relies on consumer attributes interacting with product characteristics, as in the standard BLP specification, rather than with product identities. With product-specific interactions, as in BH24's illustrative specification (their Equation (13)), the rank conditions can still hold.

[24] While the classic notion of diversion ratios focuses on price changes (e.g., Farrell and Shapiro, 2010), we work with a diversion ratio based on changes in the quality index. Price and quality diversion generally differ (Conlon and Mortimer, 2021), although they are identical in models where price only influences demand through a linear effect on the quality index (e.g., multinomial logit and nested logit).

BH24 show that the (inverse) quality substitution matrix can be identified when there are J consumer attributes that generate rich variation in shares. Intuitively, the quality substitution matrix has $J \times J$ unknown terms, so BH24's identification strategy uses $J \times J$ first-order variation: J products by J consumer attributes.[25] Identifying diversion ratios with VPA cuts down on the need for attribute-driven variation. Only the J own-quality derivatives still need to be identified, and they can be identified by $J \times 1$ first-order variation: J products by 1 consumer attribute.

Moving from intuition to formal results, we first note that invertibility (Assumption 2) implies we can express the diversion ratios as functions of the shares:

$$D_{k\to j}(s,p,x) \equiv \tilde{D}_{k\to j}\left(\sigma^{-1}(s,p,x),p,x\right).$$

We will use VPA to identify $D_{k\to j}(\cdot)$ on $\mathcal{SPX}$ for all product pairs $k \neq j$. With $D_{k\to j}(\cdot)$ identified, we can use first-order variation driven by a scalar consumer attribute to identify the inverse-demand Jacobian from (1), subject to some additional assumptions and normalizations.

Assumption 5. (Consumer-Attribute-Driven Variation) Z_{it} is scalar-valued. Furthermore, for each $(x,j) \in \mathcal{X} \times \{1,\ldots,J\}$: $g_j(\cdot,x)$ is uniformly continuous and continuously differentiable, with $\partial g_j(z,x)/\partial z \neq 0$ for all $z \in \mathcal{Z}(x)$.[26]

Before stating the normalizations, we introduce some notation. Because $g_j(\cdot)$ depends on X_t only through X_{jt} (Assumption 1), markets with different characteristic vectors are informative about the same function $g_j\left(\cdot,x_j\right)$ whenever they share the same jth coordinate. So, for each j and each x_j in the support of X_{jt}, let

$$\mathcal{Z}_j\left(x_j\right) \equiv \bigcup \left\{\mathcal{Z}(x) \;:\; x \in \mathcal{X} \text{ with } j\text{th coordinate } x_j\right\}$$

collect the attribute values at which $g_j\left(\cdot,x_j\right)$ is relevant. As a union of open sets, $\mathcal{Z}_j(x_j)$ is open, so it is a countable disjoint union of open intervals – its connected components. Each $\mathcal{Z}(x)$ is itself connected (Assumption 3), so it lies in exactly one

[25] Claim 2 gives an alternative under Assumption 3(ii), in which the $J \times J$ variation comes from higher-order derivatives of the quality indices w.r.t. a single attribute. The required order of the derivatives grows with J, so the required richness still grows with J.

[26] With Z_{it} scalar-valued, $\mathcal{Z}(x)$ connected (Assumption 3), and $\partial g_j(\cdot,x)/\partial z$ continuous and non-vanishing, $\partial g_j(\cdot,x)/\partial z$ has constant sign. So, $g_j(\cdot,x)$ is strictly monotone and $\gamma(\cdot,x,\xi)$ is injective on $\mathcal{Z}(x)$, which is BH24's Assumption 3. Requiring the condition at every z, rather than almost every z, therefore costs little and still gives us injectivity of the index.

such component, which we denote $\mathcal{C}_j(x)$. We note that this subscript indexes the product. It does not index a coordinate of z. The same is true of the reference values $z_j^0(\cdot)$ introduced below: $z_j^0(\cdot)$ is a normalization point specific to product j, not a jth consumer attribute, since Z_{it} is scalar-valued (Assumption 5).

Assumption 6. (Normalizations) For each $j \in \{1, \ldots, J\}$, each x_j in the support of X_{jt}, and each connected component $\mathcal{C}$ of $\mathcal{Z}_j\left(x_j\right)$, the researcher chooses a reference attribute value $z_j^0\left(x_j, \mathcal{C}\right) \in \mathcal{C}$ and normalizes (i) (location) $g_j\left(z_j^0\left(x_j, \mathcal{C}\right), x_j\right) = 0$; and (ii) (scale) $\partial g_j\left(z_j^0\left(x_j, \mathcal{C}\right), x_j\right) / \partial z = 1$.

Assumption 5 requires the scalar attribute Z_{it} to shift all products' quality indices. Assumption 6 imposes location and scale normalizations that apply without loss of generality.[27] We impose them separately on each connected component of $\mathcal{Z}_j(x_j)$, which is what keeps the normalized representation inside the class admitted by Assumption 1.

Together, Assumptions 1-6 lead to the first of the paper's major results: if diversion ratios are identified, then variation in shares driven by a single consumer attribute is sufficient to identify the inverse-demand Jacobian.

Proposition 1. *Suppose Assumptions 1-6 hold, and that $D_{k \to j}(s, p, x)$ is identified on $\mathcal{SPX}$ for all $k, j \in \{1, \ldots, J\}$ such that $k \neq j$. Then, $\partial \sigma^{-1}(\cdot) / \partial s$ is identified on $\mathcal{SPX}$.*

Proof. See Appendix A.1. □

The proof of Proposition 1 in Appendix A.1 adapts mechanics from BH24 to a setting where diversion ratios are already identified. To illustrate why diversion ratios help with identification, we give a sketch argument for a "large support" case where $\operatorname{supp}\{\jmath(C_{it})\} \mid (Z_{it} = z, P_t = p, X_t = x)$ is the same for all $(z, p, x) \in \mathcal{ZPX}$, so that every market share vector is achievable at every supported combination of consumer attributes, prices, and characteristics. This effectively requires a large support condition on the demand shocks; $\mathcal{E}(p, x)$ must be rich enough that a single consumer attribute value sweeps out the entirety of $\mathcal{S}(p, x) = \mathcal{S}$ as Ξ_t varies across markets. We only use large support to simplify the presentation; neither BH24's results nor ours require it.

[27] We discuss this point in more detail in Appendix B.

Fixing (p, x), we look to identify $\partial\sigma^{-1}(\cdot, p, x)/\partial s$. For illustrative purposes, we assume that the normalization point from Assumption 6 is the same for all j: $z_j^0(x_j) = z^0(x)$. For each $s \in \mathcal{S}(p, x)$, we can find some market t such that (i) $(P_t, X_t) = (p, x)$; and (ii) the share vector for attribute value $z^0(x)$ in market t is s. We then have $\bar{s}_t\left(z^0(x)\right) = s$. From the definition of inverse demand, we also have $\sigma^{-1}\left(\bar{s}_t\left(z^0(x)\right), p, x\right) = g\left(z^0(x), x\right) + \xi_t$. Taking the derivative w.r.t. z, we obtain the identity

$$\frac{\partial\sigma^{-1}(s, p, x)}{\partial s}\frac{d\bar{s}_t\left(z^0(x)\right)}{dz} = \frac{\partial g\left(z^0(x), x\right)}{\partial z}. \tag{2}$$

When diversion ratios are unknown, we can use the conditions in Claim 1. Plugging the normalization $\partial g\left(z^0(x), x\right)/\partial z = I$ into (2), we obtain

$$\frac{\partial\sigma^{-1}(s, p, x)}{\partial s} = \left[\frac{d\bar{s}_t\left(z^0(x)\right)}{dz}\right]^{-1},$$

which identifies $\partial\sigma^{-1}(s, p, x)/\partial s$ because $d\bar{s}_t\left(z^0(x)\right)/dz$ is already identified directly from the micro data. Invertibility of $d\bar{s}_t\left(z^0(x)\right)/dz$ reflects rich attribute-driven variation in market shares, requiring J consumer attributes.

Now suppose diversion ratios are already known. We have the identity $\partial\sigma^{-1}(s, p, x)/\partial s = \text{diag}(\lambda(s, p, x))^{-1}D(s, p, x)^{-1}$, where $\lambda(s, p, x)$ is the vector of own-quality derivatives: $\lambda_j(s, p, x) \equiv \partial\sigma_j\left(\sigma^{-1}(s, p, x), p, x\right)/\partial\gamma_j$. Furthermore, the scale normalization in Assumption 6 implies that $\partial g\left(z^0(x), x\right)/\partial z$ is a vector of ones. Plugging these identities into (2), some algebra yields

$$D(s, p, x)^{-1}\frac{d\bar{s}_t\left(z^0\right)}{dz} = \lambda(s, p, x).$$

This equation identifies $\lambda(s, p, x)$ because the terms on the left-hand side are already identified: the diversion ratio matrix $D(s, p, x)$ is assumed to already be identified, and $d\bar{s}_t\left(z^0\right)/dz$ is identified directly from the micro data. With $\lambda(s, p, x)$ now identified, the inverse-demand Jacobian is then identified as $\partial\sigma^{-1}(s, p, x)/\partial s = \text{diag}(\lambda(s, p, x))^{-1}D(s, p, x)^{-1}$.

Importantly, $d\bar{s}_t(\cdot)/dz$ no longer needs to have full rank; it can be a vector, thereby admitting scalar Z_{it}. The full proof of Proposition 1 – without large support – is more complicated, adapting mechanics from BH24. While the technicalities are more intricate than the large support case, the main insight is unchanged: once diversion

ratios are identified, first-order variation driven by only one consumer attribute can identify the inverse-demand Jacobian.

The BLP model from Section 2.1 helps demonstrate how much weaker this requirement is.

Remark 2. Suppose that demand is generated by the BLP model of Section 2.1, that Assumptions 3 and 5 hold, and that $D_{k\to j}(s,p,x)$ is identified on $\mathcal{SPX}$ for all $k,j \in \{1,\ldots,J\}$ such that $k \neq j$.[28] Then, Proposition 1 identifies $\partial\sigma^{-1}(\cdot)/\partial s$ on $\mathcal{SPX}$, for any J and L.

In the BLP model, with $h(\cdot)$ continuously differentiable, Assumption 5 requires $h^{(0)\prime}(z) + \sum_{\ell=1}^{L} h^{(\ell)\prime}(z)x_j^{(\ell)} \neq 0$ for each product j and each $z \in \mathcal{Z}(x)$: the consumer attribute only needs to shift each product's quality index. The condition does not become more demanding as J grows, and it involves no comparison of the derivatives across products. So, once the diversion ratios are identified, first-order variation driven by a single consumer attribute suffices to identify the inverse-demand Jacobian for demand generated by the BLP model. Whereas, neither full-menu argument from Section 2.2 applies when $J > L+1$ (Remark 1).

Our analysis in this section shows how diversion ratios – to be identified with VPA – reduce the burden on consumer-attribute-driven variation, facilitating identification in settings where attribute-driven variation alone is insufficient for non-parametric identification. The rest of the paper focuses on identifying diversion ratios with VPA.

4 Identifying Diversion Ratios Using Variation in Product Availability

In this section, we non-parametrically identify diversion ratios using various forms of VPA in micro data that supplement traditional full-menu data. In Section 4.1, we establish identification with second-choice data in the discrete choice setting. Section 4.2 considers VPA generated by randomizing the product sets offered to consumers within markets, which expands identification beyond the discrete choice setting. Section 4.2.1 covers experimental variation in the products actually offered, and Section

[28] The BLP model satisfies Assumptions 1, 2, and 4 as established in Section 2.1, and Assumption 6 is a normalization.

4.2.2 covers conjoint studies, which obtain stated choices instead of revealed choices. There, we provide sufficient conditions for identification in the presence of stated preference bias. Section 4.3 compares these three sources of VPA and discusses some practical considerations.

Remark 3. The identification results in this section (Propositions 2-4 below) do *not* require consumer-attribute-driven variation in shares.

Remark 3 highlights an important feature of the identifying arguments for diversion ratios in this section: they do not directly leverage consumer-attribute-driven variation in market shares. They can even apply when the support of the attribute(s) is a singleton. Once diversion ratios are identified, attribute-driven variation in shares is only needed to identify own-quality derivatives as in Proposition 1.

To develop intuition for our eventual identifying arguments, fix (p, x) and suppose for a moment that the researcher could guarantee γ_j is also fixed for all $j \neq k$. Then, any variation in the shares is solely due to variation in γ_k. In that case, we could identify diversion ratios using observed shares in the data, since

$$\begin{aligned}\lim_{(s_k', s_j') \to (s_k, s_j)} \frac{s_j' - s_j}{s_k' - s_k} &= \lim_{\gamma_k' \to \gamma_k} \frac{\sigma_j\left(\gamma_k', \gamma_{-k}, p, x\right) - \sigma_j\left(\gamma_k, \gamma_{-k}, p, x\right)}{\sigma_k\left(\gamma_k', \gamma_{-k}, p, x\right) - \sigma_k\left(\gamma_k, \gamma_{-k}, p, x\right)} \\ &= D_{k \to j}(s, p, x),\end{aligned}$$

where γ_{-k} is the quality vector for non-k products. We show how various types of VPA can be used to put this identification strategy into action.

4.1 Second-Choice Data

Second-choice data can be collected when consumers make discrete choices à la Section 2.1. In addition to consumers' actual purchase (first choice) when facing the full menu, second-choice data also collects information on consumers' second-most-preferred option. This comes in the form of a survey question that asks consumers who purchased each product $k \in \{1, \ldots, J\}$ which product they would have purchased (if any) had k been unavailable.[29] This is equivalent to asking *all* consumers

[29] Consumers who choose the outside (no-purchase) option are usually excluded from second-choice data. I.e., they are not asked what they would have purchased if forced to buy an inside option.

which product they would have purchased had k been unavailable, since the only consumers who would change their decision when k becomes unavailable are those who purchased it in the first place. This follows from utility maximization: anyone who did not originally purchase product k would have no incentive to change their purchase after k's removal, since their most-preferred product remains available.

Thus, second-choice data is equivalent to within-market VPA in the form of *hypothetical* scenarios where one product becomes unavailable to all consumers. That is, for each $k \in \{1, \ldots, J\}$, we effectively observe a scenario where product k is removed from the offered product set in market t without changing the value of $\{(\gamma_{jt}(z), p_{jt}, x_{jt})\}_{j \neq k}$ on $\mathcal{Z}(x_t)$, where $\gamma_{jt}(\cdot) \equiv \gamma_j(\cdot, x_{jt}, \xi_{jt})$. We denote such a hypothetical scenario as $t \setminus k$ when k is the removed product. We denote the quality index, price, and characteristic vectors for non-k products (excluding the outside option) as $\gamma_{-k,t}(\cdot)$, $p_{-k,t}$, and $x_{-k,t}$, respectively.[30] We also denote the support of γ_{-k} conditional on $(P_t = p, X_t = x)$ as $\mathcal{G}_{-k}(p, x)$, which is the projection of $\mathcal{G}(p, x)$ onto the coordinates for non-k products.

Assumption 7. Consumers' stated choices match their (counterfactual) true choices in the hypothetical scenarios posed by second-choice data.

Assumption 7 requires consumers to report their second choices truthfully. This is a substantial restriction that deserves some discussion. Note that the assumption only concerns consumers who purchased product k. For them, the second-choice question is a (stated) choice task: the menu is the actual product set in market t, at actual prices, with one product removed. In the language of marketing, this is a choice-based conjoint (Louviere and Woodworth, 1983) whose alternatives are real products. The conjoint literature finds that choice-based formats predict revealed choices reasonably well, particularly when the task resembles a decision the respondent actually faces.[31] Second-choice data is close to the best case for such a task. The respondent has recently made the original purchase, so the products, their prices, and the tradeoffs among them are fresh. Intuitively, a consumer who deliberated over the purchase is likely to have a runner-up in mind, even if lower-ranked options received

[30] Here, $\gamma_{-k,t}(z) \equiv \gamma_{-k}(z, x_{-k,t}, \xi_{-k,t})$.

[31] Green and Srinivasan (1990) and Rao (2014) review the evidence on predictive validity. Hainmueller et al. (2015) compare choice-based conjoint responses to a behavioral benchmark and find close agreement, and Kesternich et al. (2013) compare hypothetical and revealed choices in Medicare Part D. Ding et al. (2005) show that aligning respondents' incentives improves accuracy further.

less attention. So, the second-choice question asks about a comparison the consumer has already made, rather than one that must be constructed for the survey. We expect Assumption 7 to be most credible in settings with considered purchases, such as automobiles (Berry et al., 2004; Grieco et al., 2024), and less credible for habitual purchases, where the second choice may never have been considered.

Assumption 7 rules out systematic differences between stated and revealed choices. We do not relax it in this section. Section 4.2.2 shows how conjoints can identify diversion ratios in the presence of such differences, but the argument there requires product subsets to be assigned independently of consumers' decision rules. Second-choice data violates this requirement, since only consumers who purchased product k are asked about its removal.

In Section 2.2, we defined $\bar{s}_t(z)$ as the share vector for consumers with attribute value z in market t, which was identified directly from the full-menu data. Similarly, we now define $\bar{s}_{t\backslash k}(z)$ as the share vector for non-k products in the hypothetical scenario $t \setminus k$. This is directly identified by the second-choice data under Assumption 7, since $\bar{s}_{t\backslash k}(z) = E\left[Q_{i,t\backslash k} \mid Z_{it} = z;\, t\right]$ and the individual-level realizations of $Q_{i,t\backslash k}$ are either stated in the second-choice data or inferred from non-k purchases when k was available. Under Assumption 1′, we then have

$$\bar{s}_{t\backslash k}(z) = \sigma_{\backslash k}\left(\gamma_{-k,t}(z), p_{-k,t}, x_{-k,t}\right).$$

Like $\sigma(\cdot)$, we assume that $\sigma_{\backslash k}(\cdot)$ is invertible in its first argument.

Assumption 8. (Second-Choice Invertibility) For each $(p, x, k) \in \mathcal{PX} \times \{1, \ldots, J\}$, the function $\sigma_{\backslash k}(\cdot, p_{-k}, x_{-k})$ is injective on $\mathcal{G}_{-k}(p, x)$.

Assumption 8 has an important implication: given (p, x), markets with identical $\bar{s}_{t\backslash k}(z)$ have identical $\gamma_{-k,t}(z)$. Thus, we can fix quality for all non-k products by fixing the second-choice shares. We can then identify diversion ratios using variation in first-choice shares across markets with identical second-choice shares.

Formally, for each $(s, p, x) \in \mathcal{SPX}$, we can identify $D_{k\to j}(s, p, x)$ as follows. We first find a market t such that (i) $(p_t, x_t) = (p, x)$; and (ii) there is some attribute value $\tilde{z}$ with $\bar{s}_t(\tilde{z}) = s$.[32] We then find a sequence of markets $\{t^{(\ell)}\}_{\ell=1}^{\infty}$ such that (i) $\bar{s}_{t^{(\ell)}\backslash k}(\tilde{z}) = \bar{s}_{t\backslash k}(\tilde{z})$; (ii) $\bar{s}_{k,t^{(\ell)}}(\tilde{z}) > \bar{s}_{kt}(\tilde{z})$ for each $\ell \geq 1$; and (iii) $\lim_{\ell\to\infty} \bar{s}_{t^{(\ell)}}(\tilde{z}) =$

[32] The definition of $\mathcal{S}(p, x)$ ensures that this is possible.

$\bar{s}_t(\tilde{z})$. Such a sequence exists because our assumptions imply that, conditional on (z, p, x) and the second-choice shares, (a) full-menu shares vary only due to changes in γ_k; and (b) the attainable values of ξ_k form an open set, so γ_k varies locally.[33]

This sequence of markets has the same second-choice shares as market t – and therefore the same values of $\gamma_{-k}(\tilde{z})$ by Assumption 8. So, all variation in shares is driven by variation in γ_k, induced by variation in ξ_k. We then identify $D_{k\to j}(s, p, x)$ for each $j \neq k$ as

$$D_{k\to j}(s,p,x) = \lim_{\ell\to\infty} \frac{\bar{s}_{j,t^{(\ell)}}(\tilde{z}) - \bar{s}_{j,t}(\tilde{z})}{\bar{s}_{k,t^{(\ell)}}(\tilde{z}) - \bar{s}_{k,t}(\tilde{z})}.$$

Proposition 2. *Suppose Assumptions 1′, 2-4, and 7-8 hold. Then, second-choice data identifies $D_{k\to j}(s, p, x)$ on $\mathcal{SPX}$ for all $k, j \in \{1, \ldots, J\}$ such that $k \neq j$.*

Proof. Follows from the discussion preceding the proposition, since it applies to any $(s, p, x) \in \mathcal{SPX}$ and $k \neq j$. □

4.2 Randomized Product Availability

In this section, we consider two more forms of VPA, which we group under the heading of randomized product availability. First, we consider experimental variation in the product subsets actually offered to consumers within a market at the time of purchase, from field experiments where the researcher directly manipulates product availability. Second, we consider conjoint studies that ask consumers what they would consume when offered randomized product subsets that vary across consumers. In that case, we show how to obtain identification even in the presence of stated-preference bias. Importantly, both types of VPA accommodate more general demand environments than discrete choice, including demand generated by continuous choice or behavior that does not maximize utility. This stands in contrast to second-choice data, where we relied on the utility-maximizing discrete choice framework.

[33] As discussed earlier, fixing $(P_t, X_t) = (p, x)$ and the second-choice shares fixes γ_{-k}. Because $\gamma_{-k,t}(\tilde{z}) = g_{-k}(\tilde{z}, x) + \Xi_{-k}$, fixing $\tilde{z}$ also fixes $\Xi_{-k} = \xi_{-k}$. Any remaining variation in shares reflects variation in γ_k, which in turn is driven by variation in Ξ_k. With open support for $\Xi_t \mid (P_t = p, X_t = x)$ (Assumption 3), the set $\{\xi_k : (\xi_k, \xi_{-k}) \in \mathcal{E}(p, x)\}$ is also open. Additionally, $\sigma_k(\cdot, p, x)$ is locally strictly monotone in γ_k, since $\partial\sigma_k/\partial\gamma_k$ is continuous and non-zero (Assumption 4). So, we have open support for product k's shares, conditional on (z, p, x) and the second-choice shares. A sequence $\xi_k^{(\ell)} \to \xi_k$ within this set, from the side on which σ_k increases, then satisfies condition (i) because ξ_{-k} is fixed, condition (ii) by local strict monotonicity, and condition (iii) by continuity of $\sigma(\cdot, p, x)$. In turn, since $\sigma^{-1}(\cdot, p, x)$ is continuous, (iii) ensures that $\xi_k^{(\ell)} \to \xi_k$, so the ratio below converges to $D_{k\to j}(s, p, x)$.

Remark 4. The results in this section (Propositions 3 and 4 below) do not require discrete choice or utility maximization.

Before proceeding, we discuss some formal preliminaries. In addition to the micro data on quantity-attribute pairs when the full product set is available, we assume that the researcher has data on quantities – either realized or stated – for consumers offered different product subsets within markets. Formally, let $\mathcal{J}_{it}$ denote the product set offered to consumer i in market t. This can be the full product set, or it can belong to a collection of N subsets of products, which we denote $\{\mathcal{J}_n\}_{n=1}^N$ with $\mathcal{J}_n \subset \{1,\ldots,J\}$.[34] Given a product subset $\mathcal{J}$, let $P_{\mathcal{J}t}$ denote prices for those products, $X_{\mathcal{J}t}$ denote their characteristics, etc. We denote the support of $\gamma_{\mathcal{J}}$ conditional on $(P_t = p, X_t = x)$ as $\mathcal{G}_{\mathcal{J}}(p,x)$, which is the projection of $\mathcal{G}(p,x)$ onto the coordinates in $\mathcal{J}$.

Next, let $\{\mathcal{J}\}_{\setminus k} \subset \{\mathcal{J}_n\}_{n=1}^N$ denote the collection of product subsets where product k is excluded. We make an important coverage assumption on these collections of product subsets.

Assumption 9. (Rival Product Coverage) For each $k \in \{1,\ldots,J\}$, (i) $\{\mathcal{J}\}_{\setminus k}$ is non-empty; and (ii) for each $j \in \{1,\ldots,J\} \setminus \{k\}$, there is some $\mathcal{J} \in \{\mathcal{J}\}_{\setminus k}$ such that $j \in \mathcal{J}$.

Assumption 9 requires that each product k is excluded in at least one of the product sets shown to consumers. Additionally, each other product appears in some product set that excludes product k. This assumption is satisfied with single-product removals, where $N = J$ and every product subset removes a different product. We note that second-choice data effectively replicates that scenario.[35] We also note that natural experiments, like stock-out events, can have difficulty satisfying Assumption 9 because they do not let the researcher choose the product subsets. For example, low-demand products may never stock out.

With multiple product removals, whether Assumption 9 holds depends on how the subsets are designed. For example, suppose $J = 4$ and some consumers are offered one of two alternative product sets, each with two products removed: $\{1,2\}$ or $\{3,4\}$.

[34] In the discrete choice case, the outside (no-purchase) option is always available.

[35] An important distinction between second-choice data and randomized product availability is in who faces which product subset(s). With randomized availability, each consumer faces only *one* subset or the full menu. With second-choice data, it is as if each consumer faces *every* product subset, in addition to facing the full menu.

Then, the first condition of Assumption 9 is satisfied but the second is not; product 2 is not included in the subset that excludes product 1, and vice versa. Presenting all (4 choose 2) = 6 pairs of products would satisfy the condition. Although, we can achieve coverage with even fewer subsets of product pairs. For example, the four two-product sets $\{1,2\}$, $\{1,3\}$, $\{2,4\}$, and $\{3,4\}$ satisfy both conditions in Assumption 9. Every product appears in exactly two of them, and no two products appear in the same two. So, coverage is achieved with four subsets rather than six.

Importantly, Assumption 9 does not require many product subsets when they are designed carefully. The design of such subsets is a case of the well-studied "completely separating system" problem (Dickson, 1969; Spencer, 1970). Fix N such that $\binom{N}{\lfloor N/2 \rfloor} \geq J$ and assign each product a distinct label consisting of $\lfloor N/2 \rfloor$ of the indices $\{1,\ldots,N\}$.[36] Then, offer product set n to the products whose labels contain n. Each product is excluded from about half of the subsets, and every ordered pair (k,j) is separated: product j's label contains some index that product k's label does not, since the two labels are distinct and have the same size.

The requirement that $\binom{N}{\lfloor N/2 \rfloor} \geq J$ is necessary; no design with fewer subsets satisfies Assumption 9.[37] So, $N = 5$ subsets suffice for $J = 10$ products, $N = 6$ for $J = 20$, $N = 8$ for $J = 50$, and $N = 9$ for $J = 100$. In general, the minimum required N grows with $\log_2 J$ rather than with J. We note that subsets removing many products at once are more demanding in other respects, which we discuss in Section 4.3.

4.2.1 Experimental Variation

Here, we consider VPA induced by experimental variation in the product set each consumer actually faces within a market, with data on revealed quantities given the product set offered. Let $Q_{i\mathcal{J}t}$ denote the quantities consumer i demands when offered product set $\mathcal{J}$ in market t, and let $s_{|\mathcal{J}}(C_{it}) = E\left[Q_{i\mathcal{J}t} \mid C_{it}\right]$ denote the corresponding demand function. We impose an index restriction on demand for each product subset, similar to Assumption 1 for the full product set.

Assumption 10. (Index Sufficiency for Product Subsets) For each $\mathcal{J} \in \{\mathcal{J}_n\}_{n=1}^{N}$,

[36] $\lfloor N/2 \rfloor$ denotes the largest integer no greater than $N/2$; e.g., $\lfloor 5/2 \rfloor = 2$ and $\lfloor 6/2 \rfloor = 3$.

[37] Let $T_j \subseteq \{1,\ldots,N\}$ index the subsets that include product j. Assumption 9(ii) requires $T_j \not\subseteq T_k$ for all $j \neq k$, so no T_j may contain another. By Sperner's theorem, at most $\binom{N}{\lfloor N/2 \rfloor}$ such sets exist.

$\mathfrak{s}_{|\mathcal{J}}(C_{it}) = \sigma_{|\mathcal{J}}\left(\gamma_{\mathcal{J}}(Z_{it}, X_{\mathcal{J}t}, \Xi_{\mathcal{J}t}), P_{\mathcal{J}t}, X_{\mathcal{J}t}\right)$, where $\gamma_{\mathcal{J}}(\cdot)$ collects the components of $\gamma(\cdot)$ in Assumption 1 for the products in $\mathcal{J}$.

Assumption 10 is an intuitive, albeit substantial, extension of the index sufficiency condition in Assumption 1. It implicitly imposes index consistency across offered product sets: removing products does not change the quality index for the products that remain. Notably, Assumption 10 holds under Assumption 1′, which is how Section 4.1 obtains index sufficiency for single-product removals. Although, we stress that Assumption 10 is not limited to the discrete choice setting.

Next, let consumer i's *decision rule* be the random function that determines $Q_{i\mathcal{J}t}$ for each product set $\mathcal{J}$ that could be offered. We assume that product sets are assigned exogenously w.r.t. this decision rule.

Assumption 11. (Exogenous Product Sets) (i) $\mathcal{J}_{it}$ is independent of consumer i's decision rule, conditional on C_{it}; and (ii) $\Pr\left(\mathcal{J}_{it} = \mathcal{J} \mid C_{it}\right) > 0$ for $\mathcal{J} = \{1, \ldots, J\}$ and for each $\mathcal{J} \in \{\mathcal{J}_n\}_{n=1}^N$.[38]

Assumption 11(i) requires product sets to be assigned independently of consumers' decision rules, so variation in product sets is exogenous. This is a substantial restriction; e.g., it may not be satisfied by natural experiments like stock-outs, which can depend on consumers' decisions.[39] Part (ii) says that every consumer can face the full product set, and each alternative product set, with positive probability. Second-choice data satisfies both trivially, since it is equivalent to a setting where every consumer faces every (hypothetical) alternative product set as well as the full product set. Next, we assume that demand for each product subset is invertible in its quality index.

Assumption 12. (Product Subset Invertibility) For each $(p, x, \mathcal{J}) \in \mathcal{P}\mathcal{X} \times \{\mathcal{J}_n\}_{n=1}^N$, the function $\sigma_{|\mathcal{J}}(\cdot, p_{\mathcal{J}}, x_{\mathcal{J}})$ is injective on $\mathcal{G}_{\mathcal{J}}(p, x)$.

Under Assumptions 9-12, we can hold γ_{-k} fixed by fixing the shares on the alternative product sets, generalizing the identification strategy in Section 4.1. To that end, we fix shares for all the alternative product sets that exclude product k but allow shares to vary when the full product set is offered. By Assumption 9, each non-k

[38] In the discrete choice model of Section 2.1, the decision rule is generated by $U_{it}(\cdot)$ and utility maximization. Assumption 11(i) is then implied by $\mathcal{J}_{it} \perp\!\!\!\perp U_{it}(\cdot) \mid C_{it}$.

[39] Under additional assumptions, Conlon and Mortimer (2013) argue that stock-outs can also fulfill the exogeneity condition. The requisite assumptions are straightforward but may be difficult to verify, whereas experimental variation can guarantee exogeneity by design.

product is represented in at least one product subset that excludes k, so fixing shares for those subsets fixes the quality index for all non-k products. Variation in the shares when all products are offered is then driven only by variation in the quality index for k. We use this intuition to construct sequences of markets, similar to Section 4.1.

Formally, let $\bar{s}_{t|\mathcal{J}}(z) = E\left[Q_{i\mathcal{J}t} \mid Z_{it} = z, \mathcal{J}; t\right]$ be the vector of shares for attribute value z in market t when offered product set $\mathcal{J}$.[40] We have $\bar{s}_{t|\mathcal{J}}(z) = \sigma_{|\mathcal{J}}\left(\gamma_{\mathcal{J}t}(z), p_{\mathcal{J}t}, x_{\mathcal{J}t}\right)$ under Assumptions 10 and 11(i), since conditioning on the offered product set does not select on decision rules. For each $(s, p, x) \in \mathcal{SPX}$, we once again find a market t such that (i) $(p_t, x_t) = (p, x)$; and (ii) there is some attribute value $\tilde{z}$ with $\bar{s}_t(\tilde{z}) = s$. We then find a sequence of markets $\{t^{(\ell)}\}_{\ell=1}^{\infty}$ such that (i) $\bar{s}_{t^{(\ell)}|\mathcal{J}}(\tilde{z}) = \bar{s}_{t|\mathcal{J}}(\tilde{z})$ for all $\mathcal{J} \in \{\mathcal{J}\}_{\backslash k}$; (ii) $\bar{s}_{k,t^{(\ell)}}(\tilde{z}) > \bar{s}_{kt}(\tilde{z})$ for each $\ell \geq 1$; and (iii) $\lim_{\ell\to\infty} \bar{s}_{t^{(\ell)}}(\tilde{z}) = \bar{s}_t(\tilde{z})$. Variation in $\bar{s}_{t^{(\ell)}}(\tilde{z})$ is driven solely by variation in γ_k across these markets, so we can identify $D_{k\to j}(s, p, x)$ for each $j \neq k$ as

$$D_{k\to j}(s, p, x) = \lim_{\ell\to\infty} \frac{\bar{s}_{j,t^{(\ell)}}(\tilde{z}) - \bar{s}_{j,t}(\tilde{z})}{\bar{s}_{k,t^{(\ell)}}(\tilde{z}) - \bar{s}_{k,t}(\tilde{z})}.$$

Proposition 3. *Suppose Assumptions 1-4 and 9-12 hold. Then, $D_{k\to j}(s, p, x)$ is identified on $\mathcal{SPX}$ for all $k, j \in \{1, \ldots, J\}$ such that $k \neq j$.*

Proof. Follows from the discussion preceding the proposition, since it applies to every $(s, p, x) \in \mathcal{SPX}$ and $k \neq j$. □

4.2.2 Conjoint Studies

While experimental variation is sometimes available, in most settings researchers do not fully control the products offered to consumers. Neither do firms in oligopolistic markets, since each firm controls only its own products. Conjoint studies offer an alternative that is feasible almost anywhere.

A conjoint study is a survey in which each respondent completes a sequence of hypothetical choice tasks. In the standard design, each task presents a small menu of product profiles (e.g., brands, prices, and features) chosen by the researcher. The respondent then selects one profile from the menu or reports the quantity of each that they would purchase. Conjoints are cheap to run, and they are among the most

[40] $\bar{s}_t(z)$ is equivalent to $\bar{s}_{t|\mathcal{J}}(z)$ when $\mathcal{J} = \{1, \ldots, J\}$ is the full product set.

widely used tools in marketing research.[41] They have also been used in economics to elicit preferences over hypothetical alternatives (e.g., Mas and Pallais, 2017; Wiswall and Zafar, 2018; Koşar et al., 2022; Maestas et al., 2023).

We use a particular type of conjoint. The alternatives are the actual products in market t, presented with their actual prices and characteristics. So, respondents evaluate the same products as the consumers in the full-menu data, including the products' unobserved qualities Ξ_t. The only thing randomized across respondents is which subset of products appears in the menu. Each respondent's attributes Z_{it} are recorded. Respondents are drawn from market t at the time of the full-menu data, although they need not be the same consumers and they need not be purchasers.[42] Thus, this design does not exploit the ability to vary products' profiles that makes conjoints attractive elsewhere. What it exploits is that the researcher randomizes the offered product sets by construction. Respondents can also report quantities rather than a single choice, which accommodates continuous choice demand.[43]

However, there is an important concern: conjoints elicit stated choices, whereas we are interested in consumers' actual demand. Stated choices may not match revealed choices, and the literature on stated preferences has studied this issue at length.[44] We provide sufficient conditions for conjoints to identify diversion ratios in the presence of stated preference bias.

Formally, the researcher still has full-menu data that contains $(Q_{it}, Z_{it}, P_t, X_t)$ tuples for many consumers who all face the full product set, $\{1, \ldots, J\}$. The choices in the full-menu data are revealed, not stated. The researcher also has access to data from conjoint studies, where consumers are asked to state the quantities they would choose when faced with a randomized product subset, $\mathcal{J}_{it}$. We assume that the conjoints and full-menu data come from the same markets. They need not be from the same consumers, as long as conjoints are randomized as in Assumption 14 below.

Consumers facing product set $\mathcal{J}$ in the conjoint studies report stated quantities, $\hat{Q}_{i\mathcal{J}t}$, which may differ from the quantity they would actually purchase, $Q_{i\mathcal{J}t}$. Similar

[41] Green and Srinivasan (1978, 1990) review the method, Louviere and Woodworth (1983) introduce choice-based designs, and Rao (2014); Orme (2020) cover implementation. Allenby et al. (2019) discuss the economic foundations of conjoint analysis.

[42] The distribution of Z_{it} among respondents need not match its distribution in the market, since our arguments condition on z. Respondents only need to cover the attribute values of interest.

[43] Conjoint designs that elicit quantities are sometimes called "volumetric" conjoints (Kim et al., 2007).

[44] Louviere et al. (2000) review stated choice methods. See also the references in Section 4.1.

to how $Q_{i\mathcal{J}t}$ is determined by the consumer's decision rule, $\hat{Q}_{i\mathcal{J}t}$ is determined by the consumer's *statement rule* – a decision rule for stated quantities.

The stated demand function for product set $\mathcal{J}$ in market t is $\hat{\jmath}_{|\mathcal{J}}(C_{it}) = E\left[\hat{Q}_{i\mathcal{J}t} \mid C_{it}\right]$. We impose an index restriction on stated demand, similar to Assumption 10 for actual demand.

Assumption 13. (Index Sufficiency for Stated Demand) For each $\mathcal{J} \in \{\mathcal{J}_n\}_{n=1}^N$, $\hat{\jmath}_{|\mathcal{J}}(C_{it}) = \hat{\sigma}_{|\mathcal{J}}\left(\gamma_{\mathcal{J}}(Z_{it}, X_{\mathcal{J}t}, \Xi_{\mathcal{J}t}), P_{\mathcal{J}t}, X_{\mathcal{J}t}\right)$.

We note that $\hat{\sigma}_{|\mathcal{J}}(\cdot, p_{\mathcal{J}}, x_{\mathcal{J}})$ is defined on $\mathcal{G}_{\mathcal{J}}(p,x)$, the projection of $\mathcal{G}(p,x)$ onto the coordinates in $\mathcal{J}$. Assumption 13 allows the mapping from quality to stated choices to differ arbitrarily from the mapping to revealed choices. Notably, it accommodates the scale differences between stated and revealed choices used in some models combining the two (Ben-Akiva and Morikawa, 1990; Brownstone et al., 2000; Feit et al., 2010), as well as systematic over- or under-statement of purchase intentions. All are forms of stated preference bias, i.e., $\hat{\sigma}_{|\mathcal{J}}(\cdot) \neq \sigma_{|\mathcal{J}}(\cdot)$.

While Assumption 13 accommodates stated preference bias, it also imposes substantial restrictions. It rules out respondents evaluating product quality differently in the hypothetical than in the market, and it rules out any effect of the excluded products on stated choices. Intuitively, respondents may misreport what they would do, but they do not misperceive the products.

Next, we make some assumptions on exogeneity of product sets in the conjoints and invertibility of stated demand.

Assumption 14. (Exogenous Product Sets II) For consumers in the conjoint studies, (i) $\mathcal{J}_{it}$ is independent of consumer i's decision and statement rules, conditional on C_{it}; and (ii) $\Pr\left(\mathcal{J}_{it} = \mathcal{J} \mid C_{it}\right) > 0$ for each $\mathcal{J} \in \{\mathcal{J}_n\}_{n=1}^N$.

Assumption 15. (Product Subset Invertibility II) For each $(p, x, \mathcal{J}) \in \mathcal{PX} \times \{\mathcal{J}_n\}_{n=1}^N$, the function $\hat{\sigma}_{|\mathcal{J}}(\cdot, p_{\mathcal{J}}, x_{\mathcal{J}})$ is injective on $\mathcal{G}_{\mathcal{J}}(p,x)$.

Assumption 14 requires exogeneity of the product subsets presented to consumers in the conjoint, similar to Assumption 11 for direct manipulation of product availability. There are two major differences. First, Assumption 14 requires these product subsets to be (conditionally) independent of the decision rule and statement rule, while Assumption 11 only requires independence from the decision rule. Second, Assumption 14 does not require consumers in the conjoint surveys to face the full

product set (although they still could). This is because we are assuming the conjoint data is supplementing data on actual purchases from consumers in the same market who face the full product set.

The independence condition in Assumption 14 is substantial. Notably, it does *not* hold when second-choice data is used instead of conjoint analysis. This is because second-choice data has a selection aspect; it can be seen as a combination of stated and revealed preference data, since only those who purchased product k when the full product set is available are asked what their choices would be when product k is (hypothetically) removed. All other consumers' choices in that context are treated as revealed: they purchased other products anyway, so the removal of product k is assumed not to affect their choice. Independence from the decision rule is therefore violated because the consumers who purchase product k are the only ones asked to state choices when $\mathcal{J} = \{1, \ldots, k-1, k+1, \ldots, J\}$. Thus, the techniques we use in the remainder of the section cannot handle a situation where second-choice data exhibits stated choice bias.

Assumption 15 imposes an invertibility condition, similar to Assumption 12 for revealed choices. This assumption is similar to a key identifying assumption in Meango et al. (2025), and we use it for a similar purpose here: it enables us to use stated choices to control for unobserved product quality, conditional on consumer attributes, prices, and product characteristics.

With these assumptions in place, we can straightforwardly adapt the identifying argument from experimental variation (Section 4.2.1) to conjoint studies. To that end, let $\hat{s}_{t|\mathcal{J}}(z) = E\left[\hat{Q}_{i\mathcal{J}t} \mid Z_{it} = z, \mathcal{J}; t\right]$ be the "stated shares" for consumer attribute value z in market t when offered product subset $\mathcal{J}$. We have $\hat{s}_{t|\mathcal{J}}(z) = \hat{\sigma}_{|\mathcal{J}}\left(\gamma_{\mathcal{J}t}(z), p_{\mathcal{J}t}, x_{\mathcal{J}t}\right)$ by Assumptions 13 and 14(i), since the latter implies that the offered product set does not select on either decision or statement rules.

For each $(s, p, x) \in \mathcal{SPX}$, we again find a market t where (i) $(p_t, x_t) = (p, x)$; and (ii) there is an attribute value $\tilde{z}$ with $\bar{s}_t(\tilde{z}) = s$. Next, find a sequence of markets $\{t^{(\ell)}\}_{\ell=1}^{\infty}$ such that (i) $\hat{s}_{t^{(\ell)}|\mathcal{J}}(\tilde{z}) = \hat{s}_{t|\mathcal{J}}(\tilde{z})$ for all $\mathcal{J} \in \{\mathcal{J}\}_{\backslash k}$; (ii) $\bar{s}_{k,t^{(\ell)}}(\tilde{z}) > \bar{s}_{kt}(\tilde{z})$ for each $\ell \geq 1$; and (iii) $\lim_{\ell \to \infty} \bar{s}_{t^{(\ell)}}(\tilde{z}) = \bar{s}_t(\tilde{z})$. Note that (i) involves stated choices in the conjoints, while (ii) and (iii) involve revealed choices when facing the full product set.

When the coverage property from Assumption 9 holds, $\hat{s}_{t^{(\ell)}|\mathcal{J}}(\tilde{z}) = \hat{s}_{t|\mathcal{J}}(\tilde{z})$ for all $\mathcal{J} \in \{\mathcal{J}\}_{\backslash k}$ implies that $\gamma_{-k,t^{(\ell)}}(\tilde{z}) = \gamma_{-k,t}(\tilde{z})$ for every ℓ in the sequence. Thus,

all variation in chosen shares $\bar{s}_{t^{(\ell)}}(\tilde{z})$ is due to variation in $\gamma_{k,t^{(\ell)}}(\tilde{z})$, driven solely by variation in $\xi_{k,t^{(\ell)}}$. We then identify $D_{k\to j}(s,p,x)$ for each $j \neq k$ as

$$D_{k\to j}(s,p,x) = \lim_{\ell\to\infty} \frac{\bar{s}_{j,t^{(\ell)}}(\tilde{z}) - \bar{s}_{j,t}(\tilde{z})}{\bar{s}_{k,t^{(\ell)}}(\tilde{z}) - \bar{s}_{k,t}(\tilde{z})}.$$

Proposition 4. *Suppose Assumptions 1-4, 9, and 13-15 hold. Then, $D_{k\to j}(s,p,x)$ is identified on $\mathcal{SPX}$ for all $k,j \in \{1,\ldots,J\}$ such that $k \neq j$.*

Proof. Follows from the discussion preceding the proposition, since it applies to every $(s,p,x) \in \mathcal{SPX}$ and $k \neq j$. □

4.3 Summary and Discussion

Propositions 2-4 establish that three forms of VPA identify the diversion ratios $D_{k\to j}(\cdot)$ on $\mathcal{SPX}$. Table 1 summarizes how the three differ. We discuss the main differences here, along with some practical considerations.

Relationship between the three sources. The three sources share one identifying argument: fix the shares on the product subsets that exclude product k, which fixes the quality index for all other products, and ascribe the remaining variation in shares to product k's quality. They differ in how the data on those subsets is obtained. Second-choice data delivers the single-product removals $\{1,\ldots,J\} \setminus \{k\}$, and only those. We note that our results apply without change to ranked-choice data that record additional rankings beyond the second choice, since the second choice is already sufficient to identify diversion ratios. Additional rankings deliver further hypothetical product sets with two or more products removed, so it is unclear what additional identifying content they would provide in this context. Although, they may provide over-identifying restrictions that render the model's assumptions falsifiable.

Second-choice data effectively combines stated and revealed choices, since the counterfactual choice is elicited only from consumers who purchased product k and is inferred from the first choice for everyone else. Two important caveats apply: (i) the inference from first choices relies on utility-maximizing discrete choice; and (ii) the selection into who is asked is what prevents the stated-preference correction of Section 4.2.2 from applying. A conjoint that shows every respondent the single-product

TABLE 1: Three sources of variation in product availability

	Second-choice data	Experimental variation	Conjoint studies
Choices observed	Revealed (first choice) and stated (second choice)	Revealed	Revealed (full product set) and stated (subsets)
Who reports the counterfactual	Only consumers who purchased the removed product	Every consumer assigned the subset	Every respondent assigned the subset
Product sets delivered	Single-product removals	Determined by the experiment	Researcher's choice
Demand environment	Discrete choice, utility maximization	General	General
Key source-specific assumptions	Discrete choice utility (Assumption 1′); truthful second choices (Assumption 7)	Index sufficiency on subsets (Assumption 10); exogenous assignment (Assumption 11)	Index sufficiency of stated demand (Assumption 13); exogenous assignment (Assumption 14)
Invertibility	Assumption 8	Assumption 12	Assumption 15 (stated demand)
Result	Proposition 2	Proposition 3	Proposition 4

removals is the second-choice question without the selection and applies beyond discrete choice. Experimental variation removes the hypothetical element altogether, at the cost of requiring exogenous variation in what consumers are actually offered, which is only feasible in some settings.

Designing the product subsets. Section 4.2 shows that VPA from experimental variation or conjoints over relatively few subsets that each remove many products at once can identify diversion ratios. Such subsets are more demanding in other respects. Index consistency (Assumptions 10 and 13) requires that removing several of the products leaves the remaining products' quality indices unchanged, which may be a strong assumption. Invertibility (Assumptions 12 and 15) must hold on smaller product sets. And in a conjoint, a menu with several products missing is further from the true market scenario for respondents than a menu with one product missing. Researchers can trade off these considerations against the cost of additional subsets.

Sub-sample sizes. The identifying arguments treat $\bar{s}_{t|\mathcal{J}}(z)$ or $\hat{s}_{t|\mathcal{J}}(z)$ as identified from the VPA micro data for each subset $\mathcal{J}$, each market t, and each consumer attribute value z. In practice, each is a non-parametric regression on the consumers or respondents assigned to subset $\mathcal{J}$ in market t. Consumers are therefore split across the N subsets within each market, raising potential concerns about small sample sizes within subsets in practice. This can be mitigated by designing the subsets as described in Section 4.2, which keeps N relatively small. Second-choice data economizes here as well, since all consumers contribute to the data on every hypothetical single-product removal.

Cross-market variation is still required. VPA replaces consumer-attribute-driven variation in the identification of diversion ratios. It does not replace cross-market variation in Ξ_t. Each of the arguments in Sections 4.1 and 4.2 uses a sequence of markets with the same (p, x), the same shares on the subsets excluding product k, and different shares on the full product set. Such sequences exist because Ξ_t varies across markets (Assumption 3(ii)); they cannot be constructed within a single market. So, VPA in micro data from multiple markets is needed for the results in this section, and consumer-attribute-driven variation in full-menu choices in those markets is needed for Proposition 1.

5 Conclusion

This paper examines the non-parametric identifying content of variation in product availability for differentiated products demand models. VPA from multiple markets point identifies marginal diversion ratios w.r.t. product quality. As a result, consumer-attribute-driven variation is only needed to identify the own-quality derivatives, and a single attribute can suffice where existing results – without VPA – require as many consumer attributes as products. We consider three sources of VPA. Second-choice data is familiar in the literature, but its identifying content relies on utility-maximizing discrete choice and truthful reporting of second choices. Experimental variation and conjoint studies accommodate general demand, including continuous choice and behavior other than utility maximization. Conjoints do so at the cost of relying on stated choices, and we provide conditions under which stated preference bias does not obstruct identification.

Our results have several implications for empirical practice. Our arguments use VPA collected across multiple markets, since they compare markets that differ only in the unobserved quality of one product. Second-choice data suffices where it is available and the discrete choice framework applies. A conjoint over the actual products in each market provides the same identifying content with broader applicability, and a relatively small number of well-designed product subsets suffices. More generally, our results provide a non-parametric foundation for the use of VPA in demand estimation, whether it comes from second-choice surveys, experiments, or conjoints.

Our insights also point toward future research. Our analysis of conjoints uses VPA while holding products' features and prices at their actual market values. Conjoint studies in marketing typically do the opposite: they hold the product set fixed and vary product characteristics and prices across hypothetical profiles (Green and Srinivasan, 1990; Allenby et al., 2019), and Meango et al. (2025) use covariate variation in a similar way. Such variation is an alternative avenue for identification, and combining the two designs may be fruitful.

Estimation with VPA is another important direction. Estimation with second-choice data is well developed (Berry et al., 2004; Grieco et al., 2024; Conlon and Gortmaker, 2026; Conlon et al., 2026), and Conlon and Mortimer (2010, 2013) estimate demand using experimental and quasi-experimental variation. To our knowledge, there is no counterpart for conjoints in our context. We have shown that identi-

fication is possible; developing practical estimators that exploit it remains open, and we leave this to future work.

References

Abaluck, J. and Adams-Prassl, A. (2021). What do consumers consider before they choose? identification from asymmetric demand responses. *The Quarterly Journal of Economics*, 136:1611–1663.

Ackerberg, D., Benkard, C., Berry, S., and Pakes, A. (2007). Econometric tools for analyzing market outcomes. In Heckman, J. and Leamer, E., editors, *Handbook of Econometrics*, volume 6, chapter 63, pages 4171–4276. North-Holland.

Agarwal, N. and Somaini, P. J. (2026). Demand analysis under latent choice constraints. *The Review of Economic Studies*, 93(4):2215–2249.

Allenby, G. M., Hardt, N., and Rossi, P. E. (2019). Economic foundations of conjoint analysis. In Dubé, J.-P. and Rossi, P. E., editors, *Handbook of the Economics of Marketing*, volume 1, chapter 3, pages 151–192. North-Holland.

Atalay, E., Frost, E., Sorensen, A., Sullivan, C., and Zhu, W. (2026). Scalable demand and markups. *Journal of Political Economy*, 134(8):2488–2522.

Backus, M., Conlon, C., and Sinkinson, M. (2021). Common ownership and competition in the ready-to-eat cereal industry. NBER Working Paper 28350.

Barseghyan, L., Coughlin, M., Molinari, F., and Teitelbaum, J. C. (2021a). Heterogeneous choice sets and preferences. *Econometrica*, 89:2015–2048.

Barseghyan, L., Molinari, F., and Thirkettle, M. (2021b). Discrete choice under risk with limited consideration. *American Economic Review*, 111:1972–2006.

Ben-Akiva, M. and Morikawa, T. (1990). Estimation of switching models from revealed preferences and stated intentions. *Transportation Research Part A: General*, 24(6):485–495.

Bernheim, B. D., Björkegren, D., Naecker, J., and Pollmann, M. (2021). Causal inference from hypothetical evaluations. Working Paper 29616, National Bureau of Economic Research.

Berry, S. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 25(2):242–262.

Berry, S., Gandhi, A., and Haile, P. (2013). Connected substitutes and invertibility of demand. *Econometrica*, 81(5):2087–2111.

Berry, S. and Haile, P. (2014). Identification in differentiated products markets using market level data. *Econometrica*, 82(5):1749–1797.

Berry, S. and Haile, P. (2021). Foundations of demand estimation. In Ho, K., Hortacsu, A., and Lizzeri, A., editors, *Handbook of Industrial Organization*, volume 4, pages 1–62. North-Holland.

Berry, S. and Haile, P. (2024). Nonparametric identification of differentiated products demand using micro data. *Econometrica*, 92(4):1135–1162.

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

Berry, S., Levinsohn, J., and Pakes, A. (2004). Differentiated products demand systems from a combination of micro and macro data: The new car market. *Journal of Political Economy*, 112(1):68–105.

Borusyak, K., Chen, J., Hull, P., and Lei, L. (2026). Nonparametric identification of demand without exogenous product characteristics. Working Paper, arXiv:2512.23211.

Briggs, J., Caplin, A., Leth-Petersen, S., and Tonetti, C. (2026). Identification of marginal treatment effects using subjective expectations. Working Paper.

Brownstone, D., Bunch, D. S., and Train, K. (2000). Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles. *Transportation Research Part B: Methodological*, 34(5):315–338.

Calder-Wang, S. and Kim, G. H. (2026). Algorithmic pricing in multifamily rentals: Efficiency gains or price coordination. Working Paper.

Conlon, C. and Gortmaker, J. (2026). Incorporating micro data into differentiated products demand estimation with PyBLP. *Journal of Econometrics*, 257. Article 105926.

Conlon, C. and Mortimer, J. H. (2010). Effects of product availability: Experimental evidence. Working Paper 16506, National Bureau of Economic Research.

Conlon, C. and Mortimer, J. H. (2013). Demand estimation under incomplete product availability. *American Economic Journal: Microeconomics*, 5:1–30.

Conlon, C. and Mortimer, J. H. (2021). Empirical properties of diversion ratios. *The RAND Journal of Economics*, 52(4):693–726.

Conlon, C., Mortimer, J. H., and Sarkis, P. (2026). Estimating preferences and substitution patterns from second-choice data alone. Working Paper.

Conlon, C. and Rao, N. (2025). The cost of curbing externalities with market power: Alcohol regulations and tax alternatives. Accepted at Journal of Political Economy.

Dardanoni, V., Manzini, P., Mariotti, M., and Tyson, C. J. (2020). Inferring cognitive heterogeneity from aggregate choices. *Econometrica*, 88:1269–1296.

Dearing, A., Magnolfi, L., Quint, D., Sullivan, C. J., and Waldfogel, S. B. (2026). Learning firm conduct: Pass-through as a foundation for instrument relevance. Working Paper.

Dickson, T. J. (1969). On a problem concerning separating systems of a finite set. *Journal of Combinatorial Theory*, 7:191–196.

Ding, M., Grewal, R., and Liechty, J. (2005). Incentive-aligned conjoint analysis. *Journal of Marketing Research*, 42(1):67–82.

Duarte, M., Magnolfi, L., Sølvsten, M., and Sullivan, C. (2024). Testing firm conduct. *Quantitative Economics*, 15(3):571–606.

Farrell, J. and Shapiro, C. (2010). Antitrust evaluation of horizontal mergers: An economic alternative to market definition. *The B.E. Journal of Theoretical Economics*, 10(1). Article 9.

Feit, E. M., Beltramo, M. A., and Feinberg, F. M. (2010). Reality check: Combining choice experiments with market data to estimate the importance of product attributes. *Management Science*, 56(5):785–800.

Gandhi, A. and Houde, J.-F. (2020). Measuring substitution patterns in differentiated-products industries. Working Paper.

Gandhi, A. and Nevo, A. (2021). Empirical models of demand and supply in differentiated products industries. In Ho, K., Hortacsu, A., and Lizzeri, A., editors, *Handbook of Industrial Organization*, volume 4, pages 63–139. North-Holland.

Giustinelli, P. and Shapiro, M. D. (2024). SeaTE: Subjective ex ante treatment effect of health on retirement. *American Economic Journal: Applied Economics*, 16(2):278–317.

Goeree, M. S. (2008). Limited information and advertising in the U.S. personal computer industry. *Econometrica*, 76(5):1017–1074.

Goldberg, P. K. (1995). Product differentiation and oligopoly in international markets: The case of the U.S. automobile industry. *Econometrica*, 63(4):891–951.

Green, P. E. and Srinivasan, V. (1978). Conjoint analysis in consumer research: Issues and outlook. *Journal of Consumer Research*, 5(2):103–123.

Green, P. E. and Srinivasan, V. (1990). Conjoint analysis in marketing: New developments with implications for research and practice. *Journal of Marketing*, 54(4):3–19.

Grieco, P. L. E., Murry, C., Pinkse, J., and Sagl, S. (2025). Optimal estimation of discrete choice demand models with consumer and product data. Working Paper 33397, National Bureau of Economic Research.

Grieco, P. L. E., Murry, C., and Yurukoglu, A. (2024). The evolution of market power in the U.S. automobile industry. *The Quarterly Journal of Economics*, 139(2):1201–1253.

Hainmueller, J., Hangartner, D., and Yamamoto, T. (2015). Validating vignette and conjoint survey experiments against real-world behavior. *Proceedings of the National Academy of Sciences*, 112(8):2395–2400.

Hickman, W. and Mortimer, J. H. (2016). Demand estimation with availability variation. In Basker, E., editor, *Handbook on the Economics of Retailing and Distribution*, pages 306–339. Edward Elgar.

Kesternich, I., Heiss, F., McFadden, D., and Winter, J. (2013). Suit the action to the word, the word to the action: Hypothetical choices and real decisions in Medicare Part D. *Journal of Health Economics*, 32(6):1313–1324.

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

Koşar, G., Ransom, T., and van der Klaauw, W. (2022). Understanding migration aversion using elicited counterfactual choice probabilities. *Journal of Econometrics*, 231(1):123–147.

Louviere, J. J., Hensher, D. A., and Swait, J. D. (2000). *Stated Choice Methods: Analysis and Applications*. Cambridge University Press.

Louviere, J. J. and Woodworth, G. (1983). Design and analysis of simulated consumer choice or allocation experiments: An approach based on aggregate data. *Journal of Marketing Research*, 20(4):350–367.

Maestas, N., Mullen, K. J., Powell, D., von Wachter, T., and Wenger, J. B. (2023). The value of working conditions in the United States and the implications for the structure of wages. *American Economic Review*, 113(7):2007–2047.

Manski, C. F. (1990). The use of intentions data to predict behavior: A best-case analysis. *Journal of the American Statistical Association*, 85(412):934–940.

Manski, C. F. (1999). Analysis of choice expectations in incomplete scenarios. *Journal of Risk and Uncertainty*, 19(1–3):49–66.

Manzini, P. and Mariotti, M. (2014). Stochastic choice and consideration sets. *Econometrica*, 82:1153–1176.

Mas, A. and Pallais, A. (2017). Valuing alternative work arrangements. *American Economic Review*, 107(12):3722–3759.

Masatlioglu, Y., Nakajima, D., and Ozbay, E. Y. (2012). Revealed attention. *American Economic Review*, 102:2183–2205.

McFadden, D., Talvitie, A. P., and Associates (1977). Demand model estimation and validation. Special Report UCB-ITS-SR-77-9, Institute of Transportation Studies,

University of California, Berkeley, CA. Urban Travel Demand Forecasting Project, Phase I Final Report Series, Volume V.

Meango, R., Henry, M., and Mourifié, I. (2025). Combining stated and revealed preferences. Working Paper, arXiv:2507.13552.

Miller, G., de Paula, Á., and Valente, C. (2025). Subjective expectations and demand for contraception. *Journal of Econometrics*, 249. Article 105997.

Miller, N., Sheu, G., and Weinberg, M. (2021). Oligopolistic price leadership and mergers: The United States beer industry. *American Economic Review*, 111(10):3123–3159.

Miller, N. and Weinberg, M. (2017). Understanding the price effects of the MillerCoors joint venture. *Econometrica*, 85(6):1763–1791.

Miravete, E., Seim, K., and Thurk, J. (2018). Market power and the Laffer curve. *Econometrica*, 86(5):1651–1687.

Newey, W. and Powell, J. (2003). Instrumental variables estimation of nonparametric models. *Econometrica*, 71(5):1565–1578.

Norris Keiller, A., de Paula, Á., and Van Reenen, J. (2026). Production function estimation using subjective expectations data. Working Paper, arXiv:2407.07988.

Orme, B. K. (2020). *Getting Started with Conjoint Analysis: Strategies for Product Design and Pricing Research*. Research Publishers LLC, 4th edition.

Pantano, J. and Zheng, Y. (2013). Using subjective expectations data to allow for unobserved heterogeneity in Hotz-Miller estimation strategies. Working Paper, SSRN 2129303.

Petrin, A. (2002). Quantifying the benefits of new products: The case of the minivan. *Journal of Political Economy*, 110(4):705–729.

Rao, V. R. (2014). *Applied Conjoint Analysis*. Springer.

Spencer, J. (1970). Minimal completely separating systems. *Journal of Combinatorial Theory*, 8:446–447.

van der Klaauw, W. (2012). On the use of expectations data in estimating structural dynamic choice models. *Journal of Labor Economics*, 30(3):521–554.

van Mill, J. (2002). *The Infinite-Dimensional Topology of Function Spaces*. Elsevier, Amsterdam.

Wiswall, M. and Zafar, B. (2018). Preference for the workplace, investment in human capital, and gender. *The Quarterly Journal of Economics*, 133(1):457–507.

Wiswall, M. and Zafar, B. (2021). Human capital investments and expectations about career and family. *Journal of Political Economy*, 129(5):1361–1424.

Appendix

Appendix A Proofs Omitted from the Main Text

A.1 Proof of Proposition 1

We first fix notation. For $(s,p,x) \in \mathcal{SPX}$, let $D(s,p,x)$ be the $J \times J$ matrix with (j,k) element $D_{k \to j}(s,p,x)$ for $j \neq k$ and 1 on the diagonal, and let $\lambda(s,p,x)$ be the vector of own-quality derivatives, $\lambda_j(s,p,x) = \partial \sigma_j(\gamma,p,x)/\partial \gamma_j$ at $\gamma = \sigma^{-1}(s,p,x)$. By the definition of the diversion ratios,

$$\frac{\partial \sigma(\gamma,p,x)}{\partial \gamma} = D(s,p,x)\,\mathrm{diag}\left(\lambda(s,p,x)\right) \qquad \text{at} \qquad \gamma = \sigma^{-1}(s,p,x). \tag{3}$$

Assumption 4(iii) makes $\mathrm{diag}(\lambda(s,p,x))$ non-singular. So, $D(s,p,x)$ is non-singular exactly where $\partial \sigma(\gamma,p,x)/\partial \gamma$ is, and there the Inverse Function Theorem gives

$$\frac{\partial \sigma^{-1}(s,p,x)}{\partial s} = \mathrm{diag}\left(\lambda(s,p,x)\right)^{-1} D(s,p,x)^{-1}. \tag{4}$$

Since $D(\cdot)$ is identified on $\mathcal{SPX}$ by hypothesis, it suffices to identify the J scalars $\lambda(s,p,x)$ wherever $D(s,p,x)$ is non-singular.

Next, fix a market t with $(P_t, X_t, \Xi_t) = (p,x,\xi_t)$. For $z \in \mathcal{Z}(x)$, we have $\bar{s}_t(z) = \sigma\left(g(z,x) + \xi_t, p, x\right)$, which is continuously differentiable in z by Assumptions 4(i) and 5. Differentiating and using (3),

$$\frac{d\bar{s}_t(z)}{dz} = D\left(\bar{s}_t(z),p,x\right) w_t(z), \qquad \text{where} \qquad w_{jt}(z) \equiv \lambda_j\left(\bar{s}_t(z),p,x\right) \frac{\partial g_j\left(z,x_j\right)}{\partial z}. \tag{5}$$

Here, we write $g_j\left(z,x_j\right)$ to keep visible that g_j depends on X_t only through X_{jt} (Assumption 1). Examining (5), the term $d\bar{s}_t(z)/dz$ is identified directly from the micro data, and $D(\cdot)$ is identified by hypothesis. So, $w_t(z)$ is identified wherever $D\left(\bar{s}_t(z),p,x\right)$ is non-singular. We note that $w_{jt}(z) \neq 0$ by Assumptions 4(iii) and 5. Intuitively, the diversion ratios absorb the cross-product terms, leaving one scalar equation per product, each relating an identified quantity to the product of an unknown function of s and an unknown function of z.

The remainder of the argument adapts BH24 to a scalar consumer attribute.

Lemma 1 adapts their Appendix B, Lemma 2 their Lemma 2, and the proof of the proposition their Lemma 3 and Corollary 1. Once the diversion ratios are in hand, a single consumer attribute does the work that J of them do in BH24.

Lemma 1. *Suppose Assumptions 1, 3, and 5 hold, and fix $(p,x) \in \mathcal{PX}$. There is some $\delta > 0$ such that, for all $z, z' \in \mathcal{Z}(x)$ with $|z - z'| < \delta$, the set $\mathcal{E}(p,x) \cap \left\{\mathcal{E}(p,x) + g\left(z',x\right) - g(z,x)\right\}$ is non-empty and open.*

Proof. The set is open as an intersection of open sets (Assumption 3(ii)). For non-emptiness, take $\bar{\xi} \in \mathcal{E}(p,x)$ and $\epsilon > 0$ with $B\left(\bar{\xi},\epsilon\right) \subseteq \mathcal{E}(p,x)$, where $B(\cdot,\epsilon)$ denotes an open ball in $\mathbb{R}^J$. By uniform continuity of $g(\cdot,x)$ (Assumption 5), there is $\delta > 0$ such that $|z - z'| < \delta$ implies $\|d\| < 2\epsilon$, where $d \equiv g\left(z',x\right) - g(z,x)$. Then $\bar{\xi} + d/2$ and $\bar{\xi} - d/2$ both lie in $B\left(\bar{\xi},\epsilon\right)$, so $\bar{\xi} + d/2$ lies in the set. □

Any ξ in the set of Lemma 1 satisfies $g(z,x) + \xi = g\left(z',x\right) + \xi'$ with $\xi' \equiv \xi - d \in \mathcal{E}(p,x)$. So, Lemma 1 says that nearby attribute values can attain the same quality index in two different markets, and that there is an open set of ways to do so.

Lemma 2. *Suppose Assumptions 1-5 hold, and that $D_{k\to j}(\cdot)$ is identified on $\mathcal{SPX}$ for all $k \neq j$. Fix $(p,x) \in \mathcal{PX}$ and let δ be as in Lemma 1. Then, for all $z, z' \in \mathcal{Z}(x)$ with $|z - z'| < \delta$ and each j, the ratio $\left[\partial g_j\left(z,x_j\right)/\partial z\right] / \left[\partial g_j\left(z',x_j\right)/\partial z\right]$ is identified.*

Proof. Let A be the set in Lemma 1. The set $g(z,x) + A$ is a non-empty open subset of $\mathcal{G}(p,x)$, which is the support of $\gamma\left(Z_{it}, X_t, \Xi_t\right)$ conditional on $(P_t, X_t) = (p,x)$. So, it has positive probability under that conditional distribution, and by Assumption 4(ii) it contains some $\hat{\gamma}$ at which $\partial\sigma\left(\hat{\gamma},p,x\right)/\partial\gamma$ is non-singular. Write $\hat{\gamma} = g(z,x) + \xi = g\left(z',x\right) + \xi'$ with $\xi \in A$ and $\xi' = \xi - d$, and let t and t' be the markets with $(P_t, X_t, \Xi_t) = (p,x,\xi)$ and $(P_{t'}, X_{t'}, \Xi_{t'}) = \left(p,x,\xi'\right)$. By Assumption 3(i), z and z' are both in the support of Z_{it} in both markets, and

$$\bar{s}_t(z) = \sigma\left(\hat{\gamma},p,x\right) = \bar{s}_{t'}\left(z'\right) \equiv \hat{s}.$$

Since $D\left(\hat{s},p,x\right)$ is non-singular, (5) identifies $w_{jt}(z) = \lambda_j\left(\hat{s},p,x\right)\partial g_j\left(z,x_j\right)/\partial z$ and $w_{jt'}\left(z'\right) = \lambda_j\left(\hat{s},p,x\right)\partial g_j\left(z',x_j\right)/\partial z$. Both are non-zero, and the unknown $\lambda_j\left(\hat{s},p,x\right)$ is common to the two. Dividing,

$$\frac{w_{jt}(z)}{w_{jt'}\left(z'\right)} = \frac{\partial g_j\left(z,x_j\right)/\partial z}{\partial g_j\left(z',x_j\right)/\partial z}. \tag{6}$$

Although ξ and ξ' are latent, the markets t and t' can be found in the data. Among markets with $(P_t, X_t) = (p, x)$, any pair with $\bar{s}_t(z) = \bar{s}_{t'}\left(z'\right)$ and $D\left(\bar{s}_t(z), p, x\right)$ non-singular serves: by invertibility of $\sigma(\cdot, p, x)$ (Assumption 2), such a pair satisfies $g(z, x) + \xi_t = g\left(z', x\right) + \xi_{t'}$, so (6) holds for it. Both conditions are observable, since $D(\cdot)$ is identified. □

We can now prove the proposition.

Proof of Proposition 1. Step 1 (ratios at a fixed characteristic vector). Fix $x \in \mathcal{X}$, any p with $(p, x) \in \mathcal{PX}$, and δ as in Lemma 1. Since Z_{it} is scalar-valued and $\mathcal{Z}(x)$ is open and connected (Assumptions 3(i) and 5), $\mathcal{Z}(x)$ is an open interval. So, any $z, z' \in \mathcal{Z}(x)$ are linked by finitely many points $\zeta_0 = z', \zeta_1, \ldots, \zeta_n = z$ in $\mathcal{Z}(x)$ with $|\zeta_r - \zeta_{r-1}| < \delta$, and

$$\frac{\partial g_j\left(z, x_j\right)/\partial z}{\partial g_j\left(z', x_j\right)/\partial z} = \prod_{r=1}^{n} \frac{\partial g_j\left(\zeta_r, x_j\right)/\partial z}{\partial g_j\left(\zeta_{r-1}, x_j\right)/\partial z}$$

is identified by Lemma 2, for all $z, z' \in \mathcal{Z}(x)$ and each j.[45]

Step 2 (ratios across characteristic vectors, and the scale normalization). Fix j, and let z_j^0 be the reference attribute value $z_j^0\left(x_j, \mathcal{C}_j(x)\right)$ that Assumption 6 assigns to the component $\mathcal{C}_j(x)$ of $\mathcal{Z}_j\left(x_j\right)$ containing $\mathcal{Z}(x)$. Since z_j^0 and $\mathcal{Z}(x)$ lie in the same component, there is a finite chain of attribute supports $\mathcal{Z}\left(x^0\right), \ldots, \mathcal{Z}\left(x^K\right)$, each from a market whose characteristic vector has jth coordinate x_j, with $z_j^0 \in \mathcal{Z}\left(x^0\right)$, $\mathcal{Z}\left(x^K\right) = \mathcal{Z}(x)$, and consecutive supports overlapping (see, e.g., van Mill, 2002, Lemma 1.5.21). Pick $\hat{z}^k \in \mathcal{Z}\left(x^{k-1}\right) \cap \mathcal{Z}\left(x^k\right)$ for $k = 1, \ldots, K$, and set $\hat{z}^0 = z_j^0$ and $\hat{z}^{K+1} = z$ for any $z \in \mathcal{Z}(x)$. Consecutive points $\hat{z}^k, \hat{z}^{k+1}$ lie in the same support $\mathcal{Z}\left(x^k\right)$, so Step 1 identifies each factor in

$$\frac{\partial g_j\left(z, x_j\right)}{\partial z} = \prod_{k=0}^{K} \frac{\partial g_j\left(\hat{z}^{k+1}, x_j\right)/\partial z}{\partial g_j\left(\hat{z}^k, x_j\right)/\partial z},$$

where the product telescopes and the scale normalization $\partial g_j\left(z_j^0, x_j\right)/\partial z = 1$ closes it. The factors compose across markets with different characteristic vectors because every

[45] With a scalar consumer attribute the chain runs along a segment. BH24 need a more general chaining argument in their Lemma 3 because their $\mathcal{Z}(x)$ is a connected open subset of $\mathbb{R}^J$.

x^k has jth coordinate x_j, so $\partial g_j\left(\cdot, x_j\right)/\partial z$ is one function throughout (Assumption 1). Thus, $\partial g_j\left(z, x_j\right)/\partial z$ is identified for all $z \in \mathcal{Z}(x)$, all j, and all $x \in \mathcal{X}$.[46]

Step 3 (own-quality derivatives and the Jacobian). Take any $(s, p, x) \in \mathcal{SPX}$ at which $D(s, p, x)$ is non-singular, which is where $\partial\sigma^{-1}(s, p, x)/\partial s$ is defined, and which is almost every $s \in \mathcal{S}(p, x)$ by Assumption 4(ii). So, identification at these points is identification of $\partial\sigma^{-1}(\cdot, p, x)/\partial s$ wherever it exists, which is the sense in which BH24's Corollary 1 identifies it on $\mathcal{S}(p, x)$. By the definition of $\mathcal{S}(p, x)$, there are $z \in \mathcal{Z}(x)$ and a market t with $(P_t, X_t) = (p, x)$ such that $\bar{s}_t(z) = s$. By (5),

$$\lambda_j(s, p, x) = \frac{w_{jt}(z)}{\partial g_j\left(z, x_j\right)/\partial z}, \qquad j = 1, \ldots, J,$$

where the numerator is identified because $D(s, p, x)$ is non-singular, and the denominator is identified by Step 2 and non-zero by Assumption 5. Substituting into (4) then identifies $\partial\sigma^{-1}(s, p, x)/\partial s$. □

Remark 5. Beyond noting in Step 3 that $D(s, p, x)$ is non-singular for almost every $s \in \mathcal{S}(p, x)$, the proof uses the non-singularity condition of Assumption 4(ii) only through the observation, in Lemma 2, that a non-empty open subset of $\mathcal{G}(p, x)$ cannot be contained in a null set. This is what allows Lemma 2 to hold for all nearby pairs rather than almost all, so that no exceptional sets need to be tracked through the chaining in Steps 1 and 2. In particular, the proof places no condition on the reference attribute value z_j^0 beyond membership in its component.

Remark 6. The two parts of Assumption 3 play different roles. Part (i) allows the same consumer attribute value to be used in two different markets in Lemma 2, and makes $\mathcal{Z}(x)$ an interval in Step 1. Part (ii) is used in Lemma 1, which needs an open set in $\mathcal{E}(p, x)$ at one price vector for each x; since $\partial g_j\left(\cdot, x_j\right)/\partial z$ does not depend on p, this is all that Steps 1 and 2 require, and it is the requirement in BH24's Assumption 5. Part (ii) is also used in identifying $D(\cdot)$ on all of $\mathcal{SPX}$ (Propositions 2-4), which requires a local perturbation of Ξ_t at every (p, x).

Remark 7. Assumption 1 plays two roles in Step 2. Because $g_j\left(\cdot, x_j\right)$ is the same function in every market where product j has characteristics x_j, the scale normalization

[46] With the location normalization $g_j\left(z_j^0, x_j\right) = 0$ and the fundamental theorem of calculus on the interval $\mathcal{C}_j(x)$, $g_j\left(\cdot, x_j\right)$ is identified as well, as in BH24's Lemma 3. The proposition does not use this, and neither does the argument in Section 2.2, which absorbs $g_j\left(z^0(X_t), X_{jt}\right)$ into $f_j(P_t, X_t)$.

can be imposed only once per component of $\mathcal{Z}_j\left(x_j\right)$ (Appendix B). So, the reference attribute value z_j^0 need not lie in $\mathcal{Z}(x)$, and Step 2 has to move across markets with different characteristic vectors to reach it. The same property is what makes those moves free: $\partial g_j\left(\cdot, x_j\right)/\partial z$ is one function along the chain, so passing from one market to the next requires no additional data or assumption. BH24 face neither the need nor the shortcut. Their index depends on the full vector of characteristics, so their normalization is imposed market by market, at $z^0(x)$, and each x is handled on its own.

It is also useful to note what the identifying variation looks like in the data. In a market t with $(P_t, X_t) = (p, x)$, the map $z \mapsto \bar{s}_t(z)$ traces out a curve in $\mathcal{S}(p, x)$. In a second market t' with the same (p, x), it traces out the same underlying curve in the quality index, translated by $\xi_{t'} - \xi_t$. Lemma 2 uses market pairs whose curves intersect. Such intersections are generic when $\dim(Z_{it}) = J$, as in BH24. Whereas, they are not generic with a single consumer attribute. Assumption 3(ii) is what guarantees that they occur anyway, for every sufficiently small displacement along the curve $g\left(\mathcal{Z}(x), x\right)$.

A.2 Proof of Claim 2

Fix $x \in \mathcal{X}$. We first show that the condition on the derivatives implies the following spanning property: for every non-empty open interval $\mathcal{I} \subseteq \mathcal{Z}(x)$, the vectors $\{\partial g(z, x)/\partial z : z \in \mathcal{I}\}$ span $\mathbb{R}^J$. If not, there is a non-zero $a \in \mathbb{R}^J$ with $a'\,\partial g(z, x)/\partial z = 0$ for all $z \in \mathcal{I}$. Differentiating $J-1$ times gives $a'\,\partial^m g(z, x)/\partial z^m = 0$ for $m = 1, \ldots, J$ and all $z \in \mathcal{I}$. So, the $J \times J$ matrix of derivatives is singular throughout $\mathcal{I}$. Since a non-empty open subset of $\mathcal{Z}(x)$ cannot be contained in a null set, this contradicts the condition on the derivatives.

Next, fix any p with $(p, x) \in \mathcal{P}\mathcal{X}$. For $s \in \mathcal{S}(p, x)$, let $M(s)$ denote $\partial\sigma(\gamma, p, x)/\partial\gamma$ at $\gamma = \sigma^{-1}(s, p, x)$, and let

$$\mathcal{Z}_s = \left\{z \in \mathcal{Z}(x) : \sigma^{-1}(s, p, x) - g(z, x) \in \mathcal{E}(p, x)\right\}$$

be the set of attribute values that attain s in some market with $(P_t, X_t) = (p, x)$. The set $\mathcal{Z}_s$ is non-empty by the definition of $\mathcal{S}(p, x)$, and it is open because $\mathcal{E}(p, x)$ is open (Assumption 3(ii)) and $g(\cdot, x)$ is continuous. For $z \in \mathcal{Z}_s$, let t_z be the market with $\left(P_{t_z}, X_{t_z}, \Xi_{t_z}\right) = \left(p, x, \sigma^{-1}(s, p, x) - g(z, x)\right)$. By Assumptions 2 and 3(i), t_z is

the only market with $(P_t, X_t) = (p, x)$ and $\bar{s}_t(z) = s$, so it can be found in the data. Differentiating $\bar{s}_{t_z}(\cdot)$ at z,

$$v_s(z) \equiv \frac{d\bar{s}_{t_z}(z)}{dz} = M(s)\frac{\partial g(z,x)}{\partial z},$$

which is identified for every $z \in \mathcal{Z}_s$. We note that $\partial\sigma^{-1}(s,p,x)/\partial s$ exists exactly when $M(s)$ is non-singular, in which case it equals $M(s)^{-1}$.

We now construct a non-singular matrix $G(x)$ and identify $G(x)^{-1}\partial g(z,x)/\partial z$ on $\mathcal{Z}(x)$. Take $\bar{\xi} \in \mathcal{E}(p,x)$ and $\epsilon > 0$ with $B\left(\bar{\xi}, \epsilon\right) \subseteq \mathcal{E}(p,x)$, and a non-empty open interval $\mathcal{I} \subseteq \mathcal{Z}(x)$ such that $\|g(z,x) - g(z',x)\| < \epsilon/2$ for all $z, z' \in \mathcal{I}$. By the spanning property, there are $z^1, \ldots, z^J \in \mathcal{I}$ such that

$$G(x) \equiv \left[\frac{\partial g(z^1,x)}{\partial z}, \ldots, \frac{\partial g(z^J,x)}{\partial z}\right]$$

is non-singular. Every γ in the open set $g(z^1,x) + B\left(\bar{\xi}, \epsilon/2\right)$ satisfies $\gamma - g(z^i,x) \in B\left(\bar{\xi}, \epsilon\right)$ for each i, so each z^i attains $\sigma(\gamma,p,x)$. As in the proof of Lemma 2, this open set contains some $\hat{\gamma}$ at which $\partial\sigma(\hat{\gamma},p,x)/\partial\gamma$ is non-singular (Assumption 4(ii)). Let $\hat{s} = \sigma(\hat{\gamma},p,x)$. Then,

$$M(\hat{s})G(x) = \left[v_{\hat{s}}(z^1), \ldots, v_{\hat{s}}(z^J)\right]$$

is identified and non-singular. So, $h(z) \equiv G(x)^{-1}\partial g(z,x)/\partial z = \left[M(\hat{s})G(x)\right]^{-1} v_{\hat{s}}(z)$ is identified for $z \in \mathcal{Z}_{\hat{s}}$.

Next, take any $s \in \mathcal{S}(p,x)$ with $M(s)$ non-singular, and suppose h is identified on a non-empty open subset $\mathcal{O} \subseteq \mathcal{Z}_s$. By the spanning property, there are $\tilde{z}^1, \ldots, \tilde{z}^J \in \mathcal{O}$ such that $H \equiv \left[h(\tilde{z}^1), \ldots, h(\tilde{z}^J)\right]$ is non-singular. Then, $M(s)G(x) = \left[v_s(\tilde{z}^1), \ldots, v_s(\tilde{z}^J)\right] H^{-1}$ is identified, and so is $h(z) = \left[M(s)G(x)\right]^{-1} v_s(z)$ for all $z \in \mathcal{Z}_s$. Let $\mathcal{U}$ be the set of attribute values at which h is identified by repeating this step. The set $\mathcal{U}$ is non-empty and open. It is also closed in $\mathcal{Z}(x)$. To see this, take any $z \in \mathcal{Z}(x)$. The set $g(z,x) + \mathcal{E}(p,x)$ is open, so it contains some γ at which $\partial\sigma(\gamma,p,x)/\partial\gamma$ is non-singular, and $z \in \mathcal{Z}_s$ for $s = \sigma(\gamma,p,x)$. If z is a limit point of $\mathcal{U}$, then $\mathcal{Z}_s$ meets $\mathcal{U}$ in a non-empty open set, and the step above adds $\mathcal{Z}_s$ to $\mathcal{U}$. Since $\mathcal{Z}(x)$ is connected (Assumption 3(i)), $\mathcal{U} = \mathcal{Z}(x)$. Since h is identified on $\mathcal{Z}(x)$ and $g(\cdot,x)$ is J times continuously differentiable, the derivatives of h up to order $J-1$ are identified as well, with $\partial^{m-1}h(z)/\partial z^{m-1} = G(x)^{-1}\partial^m g(z,x)/\partial z^m$ for $m = 1, \ldots, J$.

At $z = z^0(x)$, the scale normalization gives

$$\left[h\left(z^0(x)\right), \frac{\partial h\left(z^0(x)\right)}{\partial z}, \ldots, \frac{\partial^{J-1} h\left(z^0(x)\right)}{\partial z^{J-1}}\right] = G(x)^{-1},$$

so $G(x)$ is identified.

Finally, take any $(s, p', x) \in \mathcal{SPX}$ at which $\partial\sigma^{-1}(s, p', x)/\partial s$ exists. Since g does not depend on prices, h and $G(x)$ are the same at p'. Applying the step above at p', with $\mathcal{O} = \mathcal{Z}_s$, identifies $M(s)G(x)$, and hence $\partial\sigma^{-1}(s, p', x)/\partial s = G(x)\left[M(s)G(x)\right]^{-1}$.

Appendix B Normalizations

In this appendix, we show that the location and scale normalizations in Assumption 6 are without loss of generality. The argument follows BH24 (Section 2.5), with one difference that arises from our product-specific index restriction.

Normalizations are needed because Assumption 1 does not pin down a unique representation of (g, σ). To see this, fix constants a_j and $b_j \neq 0$ for each j and define

$$\tilde{g}_j\left(z, x_j\right) = a_j + b_j g_j\left(z, x_j\right), \qquad \tilde{\Xi}_{jt} = b_j \Xi_{jt},$$
$$\tilde{\sigma}\left(\tilde{\gamma}, p, x\right) = \sigma\left(\gamma, p, x\right) \text{ with } \gamma_j = \frac{\tilde{\gamma}_j - a_j}{b_j}.$$

Then, $\tilde{\gamma}_j = a_j + b_j\gamma_j$, and $(\tilde{g}, \tilde{\sigma})$ generates the same demand function $\jmath(\cdot)$ as (g, σ). So, the data cannot distinguish between the two representations, and we must select one before asking whether g is identified.[47] The transformed representation satisfies Assumptions 1-5 whenever the original does, since $\gamma \mapsto \tilde{\gamma}$ is an invertible affine map.

The constants may vary across markets, but only in certain ways. First, a_j and b_j cannot vary with z within a market, since $\tilde{\sigma}(\cdot)$ must recover γ from $\tilde{\gamma}$ without knowledge of z. Second, Assumption 1 requires $\tilde{g}_j$ to be a function of $\left(z, x_j\right)$ alone. So, two markets with the same x_j and overlapping consumer attribute supports must use the same a_j and b_j; otherwise, $\tilde{g}_j$ would take two different values at the same $\left(z, x_j\right)$. Whereas, two markets with the same x_j but disjoint attribute supports can

[47] Assumption 1 rules out non-linear transformations of the index, which would break the additive separability of $g_j\left(z, x_j\right)$ and Ξ_{jt}. Transformations that mix different products' indices are ruled out by Assumptions 1′, 10, and 13.

use different constants. We note that Assumption 1 only restricts the arguments of g_j; it says nothing about how the distribution of Ξ_{jt} varies with X_t. Thus, b_j may vary with the full vector x through $\tilde{\Xi}_{jt}$.

The connected components of $\mathcal{Z}_j\left(x_j\right)$ capture these requirements. Each $\mathcal{Z}(x)$ lies in a single component, and two markets with overlapping supports lie in the same component. So, if we assign a_j and b_j separately to each component, then $\tilde{g}_j$ is a well-defined function of $\left(z, x_j\right)$: the pair $\left(z, x_j\right)$ determines the component, which determines the constants. Conversely, the constants must be common throughout a component, since any two markets in the same connected component are linked by a chain of overlapping supports (see, e.g., van Mill, 2002, Lemma 1.5.21). Thus, assigning the constants per component is the finest assignment that Assumption 1 allows.

Assumption 6 selects the representation component by component. Fix a component $\mathcal{C}$ of $\mathcal{Z}_j\left(x_j\right)$ and its reference attribute value $z_j^0 = z_j^0\left(x_j, \mathcal{C}\right)$. Setting

$$b_j = \frac{1}{\partial g_j\left(z_j^0, x_j\right)/\partial z} \qquad \text{and} \qquad a_j = -b_j\, g_j\left(z_j^0, x_j\right)$$

gives $\partial \tilde{g}_j\left(z_j^0, x_j\right)/\partial z = 1$ and $\tilde{g}_j\left(z_j^0, x_j\right) = 0$. This requires $\partial g_j\left(z_j^0, x_j\right)/\partial z \neq 0$, which holds by Assumption 5 at every $z \in \mathcal{Z}(x)$, and thus at every $z \in \mathcal{Z}_j\left(x_j\right)$. So, the normalization itself does not restrict the choice of reference attribute value. Neither does the proof of Proposition 1, which uses only that z_j^0 lies in the component. Every representation satisfying Assumption 1 thus has exactly one affine counterpart satisfying Assumption 6. The normalization restricts the representation of demand, not demand itself: neither $\jmath(\cdot)$ nor $\bar{\jmath}(\cdot)$ is affected.

Three remarks are in order. First, the normalized representation has a convenient interpretation. Since $g_j\left(z_j^0, x_j\right) = 0$, the quality index at the reference attribute value is $\gamma_j\left(z_j^0, x_j, \xi_{jt}\right) = \xi_{jt}$. So, Ξ_{jt} is measured in units of the quality index, and the scale normalization sets one unit of the index equal to one unit of z at the reference attribute value. These units are specific to the component. Thus, realizations of Ξ_{jt} from markets in different components of $\mathcal{Z}_j\left(x_j\right)$ are not directly comparable. This has no consequence for our results. Markets in different components have disjoint consumer attribute supports, and the arguments in Sections 3 and 4 never compare shocks across components. The instrumental variables step in Section 2.2 pools markets across components, but it is unaffected: $E\left[\tilde{\Xi}_t \mid W_t, X_t\right] = 0$ whenever $E\left[\Xi_t \mid W_t, X_t\right] = 0$,

since b_j depends only on X_t, and completeness of $(P_t, X_t) \mid W_t, X_t$ does not involve Ξ_t.

Second, the normalizations play no role in identifying the diversion ratios (Propositions 2-4). Each $\tilde{D}_{k\to j}(\gamma, p, x)$ is the ratio of two derivatives of $\sigma(\cdot)$ w.r.t. the same index, γ_k. A change of representation multiplies both derivatives by $1/b_k$, which cancels. So, $D_{k\to j}(s, p, x)$ is the same function of (s, p, x) in every representation, and Propositions 2-4 identify it without selecting one. A normalization is only required to identify $g(\cdot)$ – and with it $\lambda(\cdot)$ and $\partial\sigma^{-1}(\cdot)/\partial s$ – in Proposition 1.

Third, we compare our normalizations to those in BH24. BH24 allow $\tilde{\gamma} = A(x) + B(x)\gamma$ for any non-singular $B(x)$, since their index is not product-specific, and they normalize at a reference attribute value $z^0(x)$ chosen separately for each x. Assumptions 1′, 10, and 13 rule out mixing across products. So, the admissible transformations here are a subset of theirs, and fewer restrictions are needed to select a representation: two per product on each component of $\mathcal{Z}_j\left(x_j\right)$, rather than $J^2 + J$ for each x. Whereas, our normalization must be imposed across markets rather than market by market. The reference attribute value z_j^0 is shared by every market in the component, and it lies outside $\mathcal{Z}(x)$ for many of them. So, the proof of Proposition 1 has to chain across markets with different characteristic vectors to reach z_j^0. Notably, the product-specific index restriction is also what makes this chaining possible, since $\partial g_j\left(\cdot, x_j\right)/\partial z$ is one function across all markets sharing x_j.