Demand with Network Externalities: Identification and an Application to the Dating Websites Industry*

Michael Sullivan

University of British Columbia

February 13, 2026

Abstract

This paper characterizes the identifiability of demand models with network externalities. Such models are generally not identifiable with market-level data, although microdata linking consumers' decisions and characteristics permit identification under plausible conditions. Identification relies on instrumental variables reflecting across-market variation in the distribution of consumer characteristics or in the characteristics of products on offer. Guided by the identification analysis, I empirically evaluate how network externalities shape the effects of consolidation in the US dating websites industry. The results suggest that the aggregate welfare loss from monopolization, which reflects pricing effects, is attenuated by welfare gains owing to network externalities.

*Email address: michael.sullivan@sauder.ubc.ca. I thank Phil Haile, Steven Berry, Katja Seim, Mitsuru Igami, and Judy Chevalier as well as seminar participants at Yale and the Competition Bureau Canada for helpful feedback. I additionally thank the Tobin Center for Economic Policy at Yale University and the Yale School of Management for their support in acquiring data used in this project. This paper draws on research supported by the Social Sciences and Humanities Research Council.

1 Introduction

Examples abound of markets in which a key part of a product's appeal to consumers is its popularity among other consumers. Users adopt messaging platforms such as WhatsApp or Facebook Messenger in order to communicate with other users; individuals wear particular clothing brands because they are fashionable among their peers; and people use face masks in part due to social pressure generated by their widespread uptake within a community. In each case, a product becomes more attractive to a consumer as it is adopted by others.

In other markets, popularity may be a liability rather than an asset. Fishers may avoid overly crowded fishing locations, commuters may shun oversubscribed transit modes, and consumers may dislike restaurants or venues that are perceived to be too busy. Here, greater adoption by others reduces a product's appeal.

In either case, consumer demand depends directly on the behaviour of other consumers. Such interactions are typically referred to as *network externalities* or *network effects*. These forces are pervasive across many digital, social, and infrastructure markets.

Beyond their prevalence, network externalities have first-order implications for the analysis of competition and market structure. When consumer welfare depends on the scale and composition of user networks, standard welfare analyses of antitrust and regulation need not apply. For example, mergers that combine previously separate networks may increase consumer surplus by internalizing network externalities, even when they reduce product variety or raises prices. Conversely, market power may be amplified when incumbents benefit disproportionately from positive feedback effects in demand. Negative network effects also introduce nuance into questions of market structure: the desirability of entry—such as that of a new transit line or fishing site—depends not only on the trade-off between variety benefits and fixed costs, but also on the extent to which entry mitigates congestion externalities.

This paper studies the implications of network externalities for the identification of differentiated-product demand models of the type pioneered by Berry et al. (1995), commonly called BLP. While BLP-style models are a workhorse of contemporary empirical industrial organization, they are typically derived in environments wherein consumer utilities depend only on their own characteristics and product attributes, and not on equilibrium adoption decisions of other consumers. Introducing network externalities fundamentally alters this structure, as market shares enter utility directly and equilibrium demand becomes the fixed point of a system of interdependent consumer choices.

The analysis of demand in the presence of network externalities poses fundamental identification challenges. These challenges arise because changes in product characteristics affect consumer choices through multiple, intertwined channels. A change in a product characteristic has a direct effect on demand, reflecting consumers' intrinsic valuation of that characteristic. For example, a reduction in price may—holding all else equal—lead consumers to adopt the product. At the same time, the resulting change in product usage feeds back into demand through network externalities. Under positive network effects, the additional consumers attracted by a price reduction increase the product's appeal to other consumers, further amplifying demand. As a result, observed changes in sales reflect a combination of direct preference responses and indirect network-driven feedback effects.

With only market-level data on product characteristics and sales, these two channels are generally inseparable. The core problem is that product characteristics cannot be varied independently of market shares: when changes in characteristics directly affect consumer choices, they thereby indirectly affect market shares. Consequently, across-market variation in characteristics and sales is insufficient to disentangle preferences for characteristics from network externalities.

The availability of microdata linking individual consumers' choices to their characteristics improves the prospects for identification. Variation in consumer characteristics within and across markets conveys distinct information about (i) the direct effect of consumer characteristics on choice probabilities and (ii) the strength of network externalities. The direct effect is identified from within-market variation in characteristics holding the network fixed. For example, within a given city, consumers with faster internet connections may be more likely to use a dating website than otherwise similar consumers with slower connections.

Cross-market variation in the distribution of consumer characteristics, by contrast, affects product usage through two channels. First, it generates the same direct effect identified from within-market variation. Second, it produces an indirect effect operating through network externalities: when a larger share of consumers in a city has fast internet, overall usage of the dating website increases, which in turn raises the product's appeal to other consumers. Once the direct effect is identified using within-market variation, cross-market variation can be used to identify the indirect network effect under appropriate assumptions on how consumer characteristics vary across markets.

My identification argument relies on instrumental variables reflecting this logic. The first class of instruments consists of functions of market-level distributions of individual characteristics, such as the share of consumers in a city with high-speed internet access. The validity of these instruments, which are often called *Waldfogel instruments*, requires that market-level consumer characteristics do not directly affect consumer tastes conditional on the consumer's own demographics. This exclusion restriction is violated if firms adjust unobserved local product quality or advertising in response to local consumer characteristics, or if consumers' preferences depend directly on the characteristics of their neighbours.

The second class of instruments consists of characteristics of competing products within the market, as in the instruments proposed by Berry et al. (1995). These instruments shift equilibrium adoption through substitution patterns without directly affecting unobserved product quality. Under the demand structure considered in this paper, these two classes of instruments—functions of market-level consumer characteristics and characteristics of rival products—are the only sources of exogenous variation capable of separately identifying direct preference heterogeneity and network externalities.

To illustrate the application of the article's identification logic and illustrate how network externalities shape the analysis of market competition, I conduct an empirical analysis of the US dating websites industry in 2007–2008. The industry witnessed considerable consolidation in the decade following this time period, with market leader Match acquiring competitors OKCupid in 2011 and Plenty of Fish in 2015, raising natural questions about the welfare implications of dating website merges.

I study demand for dating websites using online browsing microdata that provides consumer locations, characteristics, and records of browsing dating websites. The demand model features rich consumer heterogeneity based on observable characteristics. The instrumental variables that I use in estimating demand are functions of market-specific distributions of consumer characteristics; I choose the consumer characteristics on which I base my instrumental variables to mitigate concerns that these variables will violate the exclusion restriction required for their validity. These characteristics include an indicator for whether the consumer has broadband internet and measures of consumer internet usage. The model estimates suggest that network externalities are substantial: under the estimates of my preferred specification, an inframarginal user of a site values a 10% increase in the site's usership at \$6.34/month, which is about one third of the most popular site's price. Moreover, the estimates provide evidence of age-based homophily in dating website choice.

With the estimated model in hand, I analyzes sources of variation in dating websites' market shares across geography. Although dating websites offered the same user interface across the US, their market shares varied significantly across cities. I find that this variation primarily owed to network externalities, which amplify variation in market share caused by relatively small fundamental differences in tastes across markets. This result suggests that network externalities were central to dating websites' local competitive advantages.

Last, I assessment of the welfare effects of consolidation by simulating a counterfactual regime in which `match.com` monopolizes the dating websites market. Although `match.com` raises its prices by 30.5% upon becoming a monopolist, thus reducing consumer welfare by \$5.95/consumer, this welfare loss is significantly attenuated by gains in consumer enjoyment of network externalities: monopolization consolidates the dating websites market around a single firm, offering consumers of this monopolist a larger pool of potential matches. Gains from network externalities reduce overall consumer losses from monopolization by 42%. This result suggests that failing to account for network externalities may lead researchers to misstate the welfare effects of changes in market structure.

1.1 Related literature

The article's primary contribution is to extend the identification analysis of network interactions to the canonical differentiated products demand model of Berry et al. (1995). In doing so, I draw upon insights from the literatures on peer effects and social interactions as well as those from the literature on the identification of demand systems (Berry et al. 2013, Berry and Haile 2014, Berry and Haile 2016, and especially Berry and Haile 2024).

A large body of literature considers identification of network interactions, broadly defined. This includes the literature on peer effects (Manski 1993; Graham 2008; Graham 2018; Bramoullé et al. 2009a). The term *peer effects* refers to effects of an individual's peers outcomes on the individual's outcomes. Oftentimes, peer effects models feature individual covariates whose effects on group-level outcomes are amplified by peer effects; the coefficient of amplification is often called a *social multiplier* (Becker et al. 2000; Carrell et al. 2008). Angrist (2014) shows that this social multiplier is approximately equal to the ratio of two regression coefficients: one from a regression of group-mean outcomes on group-mean covariates to a regression of individual outcomes on individual covariates. Glaeser et al. (2003) similarly characterizes the social multiplier. This finding relates to my argument that network externalities may be identified by a comparison of across-market variation in covariates that both directly and indirectly affects market shares to within-market variation that has only a direct effect. My contribution to the peer effects literature is to extend its insights to BLP-style demand models.

Bramoullé et al. (2009b) considers the identification of peer effects using data on social networks, i.e., pairwise connections between economic agents. Given the typical lack of network data in industrial organization applications, I do not consider social network data in the present article.

The literature most closely related to my work considers discrete choice models of social interactions with a finite number of agents (Brock and Durlauf 2001a; Brock and Durlauf 2001b; Brock and Durlauf 2007; Durlauf and Ioannides 2010; Aradillas-López 2021). This work examines strategic interaction in choice, the identification implications of equilibrium multiplicity, and endogenous group formation.

My paper instead adopts a framework closer to differentiated-products demand models in industrial organization and addresses a distinct set of questions. I analyze a continuum of decision-makers, as is standard in IO demand estimation, and compare two empirical environments: one in which researchers observe market-level quantities and product characteristics ("market data"), and another in which individual-level choices and characteristics ("microdata") are available.

Finally, I emphasize the role of structural product- or market-level unobservables (ξ_{jt}) that generate econometric endogeneity in BLP-style models. Unlike Brock and Durlauf (2007), who focus on cases in which such unobservables are restricted or set to zero, I impose no distributional restrictions on these alternative-specific shocks.

My paper also relates to the empirical literature on consumer choice with network externalities. Some empirical papers that operate in a setting similar to that of my paper include Timmins and Murdock (2007) (which follows the model and estimation procedure of Bayer and Timmins 2007), Bayer et al. (2004), Guiteras et al. (2025), and Allende (2019). My work is somewhat less related to empirical studies of two-sided markets with indirect network externalities, e.g., Rysman (2004) and Farronato et al. (2024).

Last, my paper relates to the literature on online dating. This literature has mostly focused on activity within a dating website, e.g., Hitsch et al. (2010) and Fong (2024). My contribution is to characterize the role of network externalities in driving choice of dating website.

2 Model

This section proposes a semi-nonparametric model of demand with network externalities extending the general demand model of Berry and Haile (2024). The section also describes a model specification that is especially relevant for empirical applications. Although I often frame the model as a discrete-choice model of demand, this framing is not necessary: consumers choice probabilities can be interpreted as quantities purchased as opposed probabilities of choosing one unit of a specific alternative.

Table 1: Summary of notation

Symbol	Description	In market data?	In microdata?
M_t	Market sizes	✓	✓
s_t	Market shares	✓	✓
x_t	Market characteristics	✓	✓
ξ_t	Unobserved product qualities		
$\mathscr{s}_{it}$	Consumer choice probabilities		✓
w_{it}	Consumer characteristics		✓
d_i	Consumer demographic group		✓

The model features markets t, each with a measure M_t continuum of consumers. Each consumer chooses between J products and an outside option. Let $\mathcal{J}$ denote the set of products *excluding* the outside option. Each consumer i belongs to one demographic group d_i among D such groups, and the measure of consumers in demographic group d in market t is M_t^d. In the dating website example, the demographic groups could be defined according to age to allow preferences for other users to depend on their ages. Consumers i choose between alternatives based on observable individual characteristics w_{it}, observable market characteristics x_t, products' unobserved demand shifters ξ_{jt}^d of products in market t that may vary across demographic groups, and within-demographic-group market shares s_{jt}^d. I generally denote $\{\zeta_{jt}\}_{j\in\mathcal{J}_t}$ by ζ_t when I have defined random variables ζ_{jt} for each $j \in \mathcal{J}_t$. I similarly use ζ_t to denote $\{\zeta_{jt}^d\}_{j\in\mathcal{J}_t, d=1,\ldots,D}$ when I have defined ζ_{jt}^d for each $j \in \mathcal{J}_t$ and $d \in \{1,\ldots,D\}$.

Note that, in the market data setting, the researcher observes only market shares s_t and market characteristics x_t. I assume that there is no observed demographic variation in the market data setting and therefore set $D = 1$ when analyzing this case. In the microdata setting, we additionally observe w_{it}, d_i, and $\mathscr{s}_{ijt}$, the last of which is consumer i's probability of choosing product j conditional on w_{it}, d_i, x_t, ξ_t, and s_t. Table 1 summarizes the notation.

Consumer choice models typically feature prices as product characteristics contained in x_t. Additionally, price is usually thought to be endogenous in the sense that it is dependent on ξ_t. In what follows, I ignore price endogeneity (and, more generally, endogenous product characteristics other than s_t) to focus attention on network externalities.

The primative structural object in the model is the choice probability function σ, which provides consumer i's choice probability $\mathcal{s}_{ijt}$ for each product j as a function of various market and individual characteristics:

$$\mathcal{s}_{ijt} = \sigma_j(x_t, s_t, \xi_t, w_{it}, d_i), \tag{1}$$

Integrating over w_{it} yields the average choice probabilities function $\bar{\sigma}_t$, which is the structural object of interest when studying identification with market data: for each j,

$$\bar{\sigma}_{jt}(x_t, s_t, \xi_t) = \mathbb{E}\left[\sigma_j(x_t, s_t, \xi_t, w_{it}, d_i) \mid x_t, s_t, \xi_t, t\right].$$

The conditioning on t and the t subscript on $\bar{\sigma}_{jt}$ reflect that the distribution of consumer characteristics (w_{it}, d_i) may vary across markets t. In what follows, I use F_t^w to denote the distribution of $\{w_{ijt}\}_{j\in\mathcal{J}_t}$ in t.

To build intuition, I describe a semi-parametric specification of the general model. In this specification, consumer i selects the alternative j that maximizes the indirect utility

$$u_j(x_t, s_t, \xi_t, w_{it}, d_i, \varepsilon_{it}) = \begin{cases} x_{jt}'\beta + f_j^{d_i}(s_t) + \xi_{jt}^{d_i} + w_{it}'\lambda_j + \varepsilon_{ijt}, & j \neq 0 \\ \varepsilon_{i0t}, & j = 0. \end{cases} \tag{2}$$

where ε_{ijt} are unobservables and f_j^d are functions unknown to the econometrician. One simple specification of f_j^d in the case without multiple demographic groups is $f_j(s_t) = \gamma s_{jt}$. In this case, $\gamma > 0$ implies that consumers enjoy choosing the same product as others whereas $\gamma < 0$ implies that consumers dislike when others choose the same product.

The choice probability function in this model is, for each j,

$$\sigma_j(x_t, \xi_t, s_t, w_{it}, d_i) = \Pr\left(j = \arg\max_{k\in\{0,1,\ldots,J\}} u_{ikt}(x_t, s_t, \xi_t, w_{it}, d_i, \varepsilon_{it}) \mid x_t, \xi_t, s_t, w_{it}, d_i\right).$$

Although the model features a choice of one alternative, it can accommodate multi-homing, i.e., consumers adoptiong multiple platforms. To see how, suppose that a consumer chooses whether to join each of two platforms, and that the consumer receives utility proportional to the share of consumers available through these platforms. Assume for simplicity that there is one demographic group. In this case, we can let $j = 3$ denote the option of joining both platforms and set $f_1(s_t) = \gamma \times (s_{1t} + s_{3t})$, $f_2(s_t) = \gamma \times (s_{2t} + s_{3t})$, and $f_3(s_t) =$

$\gamma \times (s_{1t} + s_{2t} + s_{3t})$, where γ is the factor governing the strength of network externalities.

In a discrete choice model of demand without network externalities, we obtain market shares simply by integrating choice probabilities across consumers. With network externalities, market shares are instead fixed points of the function $\bar{\sigma}_t$ introduced above. For simplicity of exposition, I now focus on the $D = 1$ case; the generalization to $D > 1$ as described by Appendix B.4 is straightforward. The exogenous objects characterizing a market are the distribution of demographic characteristics F_t^w, the observed market characteristics x_t, and the unobserved product qualities ξ_t. Let $\chi_t = \{x_t, \xi_t\}$ denote both the observed and unobserved product characteristics. The endogenous variables are the market shares, which I assume to be a solution of the equation

$$s_t = \bar{\sigma}_t(x_t, \xi_t, s_t). \tag{3}$$

In the model outlined above, products' market shares rather than their total quantities (i.e., their shares times the market size) enter the choice probability functions. I later consider models in which choice probability functions depend on total quantities.

2.1 Multiple equilibria

Equilibrium market shares, i.e., solutions of (3), exist by Brouwer's fixed-point theorem when $\bar{\sigma}_t$ is continuous in s_t. The equilibrium, however, may not be unique. This complicates counterfactual analysis because the effect of changing market characteristics depends on which equilibrium is selected after the change. I aim to avoid this complication by appealing to local uniqueness of equilibria. Suppose we observe market characteristics $\chi_t = (x_t, \xi_t)$ and market shares s_t. By the implicit function theorem, there is a unique function s^* defined on a neighbourhood X of χ_t such that $s^*(\chi_t) = s_t$ and $\chi \in X$,

$$s^*(\bar{\chi}) - \bar{\sigma}_t(\bar{\chi}, s^*(\bar{\chi})) = 0$$

as long as the following matrix is nonsingular:

$$I - D_s \bar{\sigma}_t(\chi_t, s_t), \tag{4}$$

where $D_s \bar{\sigma}_t(\chi_t, s_t)$ is the derivative of $\bar{\sigma}_t(\chi_t, s_t)$ with respect to s_t. I call the unique s^* function defined in the neighbourhood of a particular χ_t an *equilibrium surface* around (χ_t, s_t). Local uniqueness of the equilibrium s_t at market characteristics χ_t means that there is a unique equilibrium surface defined around χ_t. In Appendix A, I discuss the condition that (4) must be nonsingular, and I argue that nonsingularity occurs only in knife-edge cases. As argued by Bayer et al. (2004), the concept of local uniqueness is useful because it allows for coherent statements about the effects of marginal changes in market characteristics.

2.2 Microfoundation for network externalities in the dating industry

Network externalities in demand for dating websites can be microfounded by a search-and-matching model. Online Appendix A proposes such a model that gives rise to a dependency between the number of users on a platform and a user's valuation of the platform. This model is a variant of Smith (2006)'s model of search and matching in the marriage market featuring quadratic search technology, exogenous matches and match destruction occurring at exponential rates, and users who idiosyncratically value matches with other users.

3 Identification

Berry and Haile (2014) establish that BLP-style demand models are identified with market data alone under an appropriate index assumption, the availability of suitable instrumental variables, and other reasonable assumptions. Adding network externalities to these models complicates the picture. This section begins by discussing the identification problems and their solutions in the context of a simple parametric model before generalizing this discussion to the semi-nonparametric model proposed in the preceding section.

Before turning to the identification analysis, I clarify the object of interest. My goal is to identify the average choice probability function $\bar{\sigma}_t$ in the market-level data and the choice probability function σ in the microdata. These functions map market characteristics and market shares into demanded quantities. An alternative object of identification is the reduced-form equilibrium mapping from market characteristics $\{x_t, \xi_t\}$ to market shares s_t, implicitly defined by the fixed-point condition in (3).

In models with network effects, this distinction matters. The equilibrium mapping $\{x_t, \xi_t\} \mapsto s_t$ may be identified even when the underlying choice function $\bar{\sigma}$ is not. In that case, the overall effect of x_t on equilibrium demand is pinned down, but its decomposition into intrinsic preferences for x_t and feedback through network externalities is not. Observationally equivalent models can generate the same equilibrium shares while differing in the strength of network spillovers. Identifying only the reduced-form mapping therefore leaves network externalities undetermined.

My analysis seeks to identify $\bar{\sigma}$ rather than merely the equilibrium mapping for two main reasons. First, welfare impacts of changes in the market environment depend on the strength of network externalities. For example, consider a reduction in a platform's membership price targeted to a subset of consumers. Without network externalities, ineligible consumers are unaffected. With positive network externalities, however, inframarginal ineligible consumers benefit because the price reduction expands the user base.

Second, many counterfactuals move the system outside the support of observed market structures. Platform integration, entry, and exit change equilibrium shares in ways that depend on the strength of network feedback. A reduced-form equilibrium mapping from product characteristics to market shares is not invariant to changes in market structure, and thus cannot be used to analyze counterfactuals in which market structure is altered. In assessing such counterfactuals, identification of the average choice probability function is crucial.

3.1 Illustrative model

Consider a market t in which consumers choose whether to purchase a single product. Let $y_{it} = \mathbb{1}\{u_{it} \geq 0\}$ denote an indicator for whether consumer i buys the product and specify the indirect utility u_{it} as

$$u_{it} = x_t + \gamma s_t + \xi_t - \varepsilon_{it}.$$

Here, x_t is an exogenous characteristic of the product, s_t is the product's market share, and ξ_t is an unobserved demand shifter in market t. Additionally, ε_{it} is consumer i's idiosyncratic taste for the product; assume that ε_{it} is iid according to the strictly increasing and continuous distribution function $F : \mathbb{R} \to [0, 1]$. We fix the coefficient of x_t at one as a scale normalization. The purchase probability conditional on s_t, x_t, and ξ_t is

$$\int \mathbb{1}\{x_t + \gamma s_t + \xi_t - \varepsilon_{it} \geq 0\} dF(\varepsilon_{it}) = F(x_t + \gamma s_t + \xi_t).$$

Imposing that the market share equals the consumer's purchase probability, we obtain

$$s_t = F(x_t + \gamma s_t + \xi_t)$$

or, inverting the function F,

$$F^{-1}(s_t) = x_t + \gamma s_t + \xi_t. \tag{5}$$

Suppose that we observe only s_t and x_t for a population of markets t; i.e., we are in the market-data setting. The true model primitives $\theta = (F, \gamma)$ are observationally equivalent to the alternative model primitives $\tilde{\theta} = (G_\delta, \gamma + \delta)$ for any $\delta \geq 0$ when

$$G_\delta^{-1}(s) = F^{-1}(s) + \delta s.$$

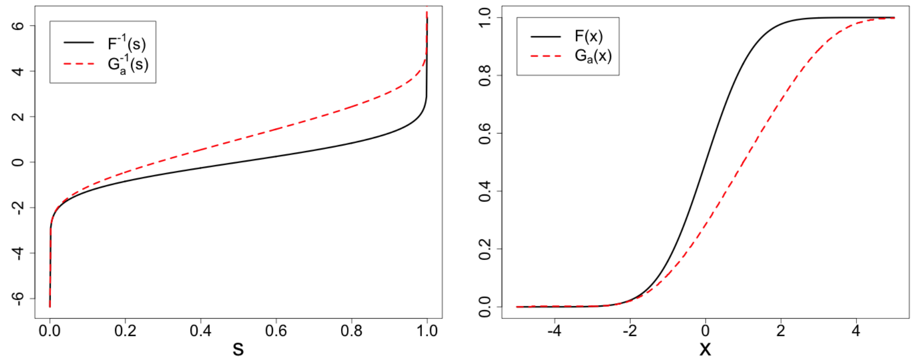
As Figure 1 illustrates, increasing δ raises the dispersion of G_δ.[1] This dampens the effect of x_t on s_t. This observational equivalence result reflects that the data cannot distinguish between the following two explanations for the observed relationship between x_t and s_t:

(i) Sales rise when x_t falls solely because consumers dislike the characteristic x_t (direct effect of x_t on market shares; this effect is relatively large when δ is low), and

(ii) Sales also rise in part because consumers enjoy the increase in users induced by the decrease in x_t (indirect effect of x_t on market shares owing to network externalities;

[1] It is straightforward to verify that G_δ is a valid distribution function. Because F is strictly increasing, its quantile function F^{-1} is strictly increasing on $(0, 1)$. For $\delta \geq 0$, the function $G_\delta^{-1}(s) = F^{-1}(s) + \delta s$ is strictly increasing and therefore admits a well-defined and strictly increasing inverse G_δ. Moreover, $\lim_{s\to 0} F^{-1}(s) = -\infty$ and $\lim_{s\to 1} F^{-1}(s) = \infty$. Since δs is bounded on $(0, 1)$, the same limits hold for $G_\delta^{-1}(s)$, implying $\lim_{x\to -\infty} G_\delta(x) = 0$ and $\lim_{x\to +\infty} G_\delta(x) = 1$. Right-continuity of G_δ follows from the continuity of F. Hence G_δ is a valid cumulative distribution function.

Figure 1: Illustration of G_δ



this effect is relatively large when δ is high).

When the product in question is a dating website, for example, a price reduction may increase adoption both because consumers are price sensitive and because the lower price expands the pool of potential matches. Incorrectly ruling out network externalities would lead the researcher to overstate the direct effect of x_t, attributing the entire relationship between x_t and s_t to tastes for x_t rather than to network externalities.

The identification problem arises because x_t affects equilibrium shares s_t, so market-level data never reveal how consumers respond to x_t, holding network size fixed. Microdata including individual consumer choices and characteristics help solve this problem. To see why, suppose that the individual characteristic w_{it} that we observe is a component of the idiosyncratic taste term ε_{it}: $\varepsilon_{it} = w_{it} + \tilde{\varepsilon}_{it}$. In the dating website setting, the consumer characteristic w_{it} could be an inverse measure of internet speed. Suppose additionally that the remaining unobservable aspect of idiosyncratic tastes $\tilde{\varepsilon}_{it}$ is distributed according to $\tilde{F}$ independently of w_{it} so that

$$u_{it} = \underbrace{\beta x_t + \gamma s_t + \xi_t}_{=:\delta_t} - w_{it} - \tilde{\varepsilon}_{it}.$$

Here, δ_t is a market-level utility index common to all consumers in market t. We now see that, when y_{it} is an indicator for whether consumer i uses the dating website and the t subscript on $\Pr_t$ indicates conditioning on all characteristics of market t,

$$\begin{aligned} \Pr_t(y_{it} = 1 \mid w_{it} = \bar{w}) &= \Pr_t(u_{it} \geq 0 \mid w_{it} = \bar{w}) \\ &= \Pr_t(\tilde{\varepsilon}_{it} \leq \delta_t - \bar{w}) \\ &= \tilde{F}(\delta_t - \bar{w}). \end{aligned} \tag{6}$$

The left-hand side is observable and the final expression on the right-hand side is the distribution function of $\delta_t - \varepsilon_{it}$. When w_{it} has large support, variation in w_{it} traces out the function $z \mapsto \tilde{F}(z)$ up to a location shift. With a location normalization (e.g. $\mathbb{E}_t[\tilde{\varepsilon}_{it}] = 0$), knowledge of this distribution separately identifies δ_t and $\tilde{F}$.

The remaining task of identification is the separate identification of the components of utility index δ_t. The primary challenge in completing this task is the fact that ξ_t and s_t are mechanically dependent because ξ_t partly determines s_t. Cross-market variation in the distribution of w_{it} shifts equilibrium shares through its effect on individual choice probabilities. If this variation is mean-independent of ξ_t (i.e., $\mathbb{E}[\xi_t \mid F_t^w] = 0$), it provides excluded instruments for s_t. [2] It follows that, as long as x_t is mean-independent of ξ_t (i.e., $\mathbb{E}[\xi_t | x_t] = 0$), each of β, γ, and ξ_t will be identified by an instrumental variables argument. Instruments based on market-specific demographic distributions are often called *Waldfogel instruments*, following the insights of Waldfogel (2003) and Waldfogel (2008) that local demographic composition shapes product availability and choice sets (Berry and Haile 2016).

The argument of Berry and Haile (2024) can be applied to identify this simple model even without a large support assumption. Their approach involves the assumption of a common choice probability s^* that can be achieved in any market t with the appropriate choice of w_{it}. In my setting, the common choice probability condition requires that there is a s^* such that for all t, the support of w_{it} includes a point $w_t(s^*)$ satisfying

$$\tilde{F}(\delta_t - w_t(s^*)) = s^*.$$

As long as $\tilde{F}$ is injective, it can be inverted in the equation above to obtain

$$w_t(s^*) = \beta x_t + \gamma s_t - \tilde{F}^{-1}(s^*) + \xi_t. \tag{7}$$

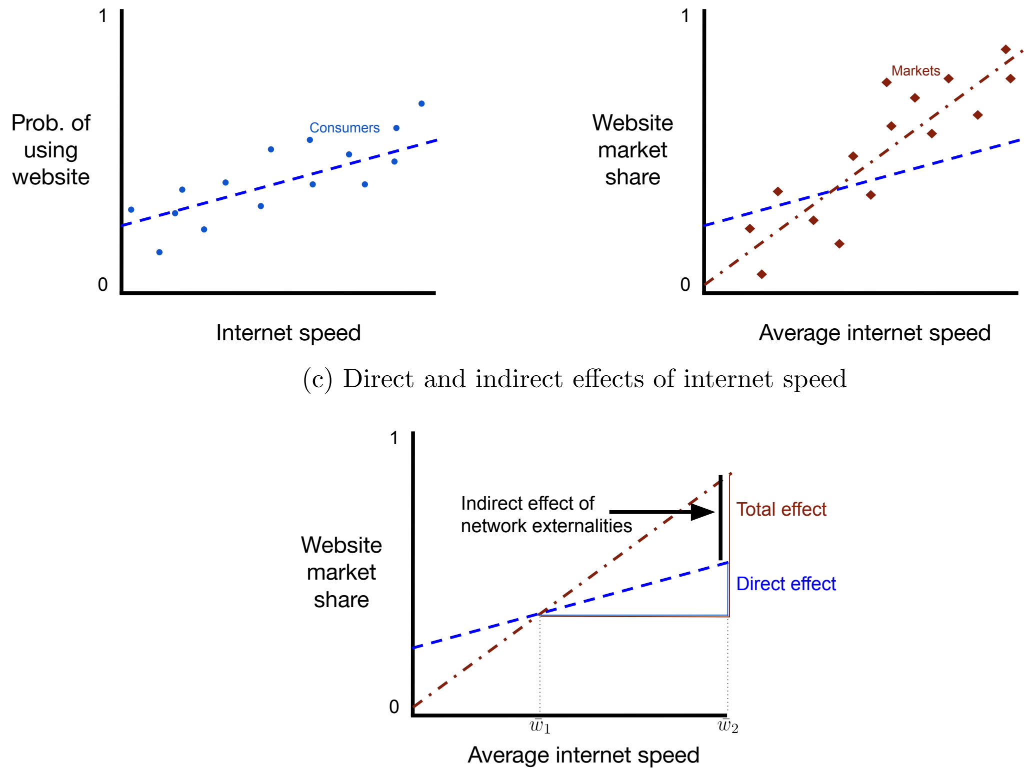
Given that the value $w_t(s^*)$ that achieves s^* in market t is observable, equation (7) is a standard regression equation with s_t as an endogenous regressor. The preceding paragraph's discussion of instruments equally applies to (7). Instruments permit the identification of ξ_t, which immediately yields the identification of the function yielding consumer choice probabilities: $(w_{it}, x_t, s_t, \xi_t) \mapsto \mathbb{E}[y_{it} | w_{it}, x_t, s_t, \xi_t]$. We obtain this function, which is typically the main structural object of interest, without necessarily point identifying $\tilde{F}$.

I now provide an example to build intuition for the identification strategy. Suppose that, within some market t, consumers with faster internet speeds are more likely to use the dating website. Figure 2a, in which consumers are represented by blue circles, expresses this relationship. Based on the observed within-market relationship between internet speed and website usage, we can predict how dating website usage will increase in t when we increase internet speeds for all consumers in the market. If there are positive network externalities (i.e., $\gamma > 0$), then this will be an underprediction: the prediction will capture the direct effect of increasing internet speeds on consumers' website usage as identified using within-market variation, but it will not capture the fact that increasing usage rates across the market increases the site's appeal due to network externalities. Figure 2b represents this scenario; in the figure, each red diamond is a market. Last, as illustrated by Figure 2c, the gap between the predicted and actual change in adoption identifies the strength of network

[2] This logic is suggested in Section 7.1.2 of Jullien et al. (2021).

Figure 2: Illustration of identification with microdata

(a) Within-market variation (b) Across-market variation



externalities.

To connect the example above to my formal identification argument, within-market variation identifies the relationship between w_{it} (which is negative internet speed in the example) and choice probabilities holding market shares fixed as shown by (6). Increasing average internet speeds shifts the distribution of w_{it} across markets. This cross-market variation identifies how equilibrium shares feed back into δ_t.

There are several plausible threats to the identification strategy using Waldfogel instruments as described above. First, the exclusion restriction may fail if consumers sort across markets based on unobserved demand shifters. For example, consumers with a strong latent taste for the dating website may disproportionately locate in areas with faster internet infrastructure. If internet speed is correlated with location and location choice depends on unobserved tastes for the website, then the cross-market distribution F_t^w of w_{it} will be correlated with ξ_t.

Second, the platform itself may respond endogenously to cross-market variation in F_t^w. If the dating website provides higher unobserved service quality or more intensive advertising in markets where the distribution of individual characteristics makes adoption especially attractive, then ξ_t becomes a function of F_t^w. Such targeted investment again violates the exclusion restriction required for instrumental variables identification.

A related concern arises from *contextual network effects*, a term that refers to situations in which an agent's outcome depends directly on the characteristics of others in the relevant group. If the distribution of characteristics directly shifts utility beyond its effect on equilibrium shares, then ξ_t necessarily depends on F_t^w, invalidating instruments constructed from that distribution. These concerns imply that the researcher must carefully select individual characteristics w_{it} whose cross-market variation is plausibly orthogonal to unobserved demand shifters and unlikely to generate endogenous firm responses.

In the binary choice setting considered above, cross-market variation in the distribution F_t^w provides the natural excluded source of variation in equilibrium shares s_t. In models with multiple alternatives, additional excluded shifters may be available. In particular, the literature on differentiated-products demand following Berry et al. (1995) uses characteristics of competing alternatives—so-called BLP instruments—to shift equilibrium shares under standard exclusion restrictions. These instruments rely on cross-market variation in the characteristics of competing products. When the set of alternatives and their observable characteristics are essentially identical across markets, however—as is often the case for platforms operating in multiple geographic regions—such variation is absent, and BLP-style instruments are unavailable. The identification strategy must then rely more heavily on variation in individual characteristics.

Although microdata are necessary to identify models in which market shares s_t directly enter individual utilities, market-level data may suffice in environments where consumers care about the total number of users rather than the share. To illustrate, suppose utility is instead given by

$$u_{it} = x_t + \gamma M_t s_t + \xi_t - \varepsilon_{it},$$

where M_t denotes the population of market t. In this formulation, network effects operate through the total number of users $M_t s_t$. Absent network externalities ($\gamma = 0$) and assuming $\mathbb{E}[\xi_t \mid M_t, x_t] = 0$, population size M_t has no direct effect on market shares conditional on observables. When $\gamma \neq 0$, however, cross-market variation in M_t shifts equilibrium adoption through its effect on the total user base. Conditional on x_t and under the exclusion assumption that M_t does not independently shift ξ_t, the correlation between population and market share identifies the sign and magnitude of γ. This approach relies on the strong restriction that population affects utility only through the network term $M_t s_t$. The restriction fails, for example, if larger markets systematically receive greater unobserved platform investment, so that ξ_t increases with M_t. In such cases, population size no longer provides valid identifying variation.

3.2 Identification with market data

I now formalize the identification arguments developed above. I consider identification of a demand system using market-level data under an index structure closely related to that of

Berry and Haile (2014):

$$\begin{aligned}\bar{\sigma}(x_t, s_t, \xi_t) &= \bar{\sigma}(\delta(x_t, s_t, \xi_t)) \\ \delta_j(x_t, s_t, \xi_t) &= x_{jt} + h_j(s_t) + \xi_{jt},\end{aligned} \tag{8}$$

where $x_t = [x_{1t}, \ldots, x_{Jt}]'$ is a vector of observable product characteristics and s_t denotes the vector of market shares.

I focus on the *average choice probability function* $\bar{\sigma}$ as the object of identification with market data. Consumer-level heterogeneity enters only through aggregation, and individual characteristics are not observed in this setting. Although $\bar{\sigma}$ may in principle depend on the distribution of consumer characteristics, such dependence cannot be exploited with market data and is therefore suppressed.

Relative to Berry and Haile (2014), the index restriction (8) imposes additional structure: endogenous market shares enter demand only through an additive index. I use a more stringent restriction to show that the model is generally not identified even in the favourable case in which market shares enter only through an additive index.[3] Note that the imposition of a coefficient of one on x_{jt} is a scale normalization. As a location normalization, I also impose $\mathbb{E}[\xi_{jt}] = 0$ for each j. I slightly abuse notation in (8) by writing $\bar{\sigma}$ both as a function of $\{x_t, s_t, \xi_t\}$ and as of the index alone. In practice, I seek to identify $\bar{\sigma}$ as a function of the index and to identify h_j, which are sufficient for identifying $\bar{\sigma}_j$ as a function of $\{x_t, s_t, \xi_t\}$.

I now state and motivate an assumption under which I study identification.

Assumption INVERT-MARKET (Invertible demand — market data)**.** The function $\bar{\sigma}$ is injective on the support of $\delta(x_t, s_t, \xi_t)$.

This assumption implies that $\bar{\sigma}$ admits a well-defined inverse $\bar{\sigma}^{-1}$ on the relevant support. Sufficient conditions for invertibility are provided in Berry et al. (2013), which require a minimal degree of substitutability across products.

The following proposition establishes that the primitives $(\bar{\sigma}, h)$ are not identified with market data.

Proposition 1. *Fix some model primitives $\theta \coloneqq (\bar{\sigma}, h)$. For any function σ satisfying Assumption INVERT-MARKET, there is a function $\tilde{h}$ and such that $\tilde{\theta} = (\tilde{\sigma}, \tilde{h})$ is observationally equivalent to θ.*

Proof. The primitives $(\bar{\sigma}, h)$ are consistent with the observable data if and only if

$$\bar{\sigma}^{-1}(s_t) = x_t + h(s_t) + \xi_t. \tag{9}$$

[3] Note also that my approach allows for product characteristics other than x_t, which are required to satisfy the index restriction stated in (8); as Berry and Haile (2014) suggest, the researcher can condition on exogenous characteristics $x_t^{(2)}$ and suppress them in the notation, implicitly identifying demand conditional on each $x_t^{(2)}$ in these characteristics' support under which the assumptions invoked to identify demand hold.

Define $\tilde{h}(s_t) = h(s_t) + \tilde{\sigma}^{-1}(s_t) - \bar{\sigma}^{-1}(s_t)$. Then,

$$\tilde{\sigma}^{-1}(s_t) = x_t + \tilde{h}(s_t) + \xi_t.$$

Thus, $\tilde{\theta} = (\tilde{\sigma}, \tilde{h})$ is observationally equivalent to θ. $\square$

Proposition 1 shows that substitution patterns (encoded in $\bar{\sigma}$) and network externalities (encoded in h) cannot be disentangled using market data alone. Market shares enter the inverse demand equation both through $\bar{\sigma}^{-1}$ and through h. Changes in equilibrium shares can therefore be rationalized either by differences in substitution patterns or by differences in the strength of network externalities. This non-identification persists even if s_t were exogenous, which is generally implausible. Note that the identification problem does not relate to the unavailability of instruments: even if s_t were mean-independent of ξ_t, which is generally impossible, the model's primitives would be unidentified.

Under the availability of instruments satisfying an appropriate exclusion restriction and completeness condition, ξ_t is identified as the residual from the nonparametric regression equation $x_{jt} = \bar{\sigma}^{-1}(s_t) - h_j(s_t) - \xi_{jt}$, where $\bar{\sigma}^{-1}(s_t) - h_j(s_t)$ is the corresponding nonparametric regression function. Identification of ξ_t implies identification of the reduced-form mapping $(x_t, \xi_t) \mapsto s_t$. However, it does not identify $\bar{\sigma}$ and h separately. Thus, even under ideal instrumental variables conditions, the structural primitives remain unidentified.

Identification problems arise even when $\bar{\sigma}$ is known, as would occur under a parametric assumption on the distribution of unobserved heterogeneity. In this case, the inverse demand equation (9) becomes

$$\bar{\sigma}_j^{-1}(s_t) = \beta_j x_{jt} + h_j(s_t) + \xi_{jt}, \quad j \in \mathcal{J} \tag{10}$$

I introduce the coefficients β_j because setting the coefficient of each x_{jt} to one is no longer a scale normalization when $\bar{\sigma}$ is known. Equation (10) contains J endogenous regressors (the elements of s_t). Standard BLP-style instruments provide at most $J-1$ excluded shifters in typical applications. Unless additional restrictions are imposed, the model remains underidentified. Parametric assumptions on substitution patterns therefore do not, by themselves, resolve the identification problem created by network externalities.

Appendix B.1 provides analysis of identification with market data in a case in which each product's δ_{jt} index depends only on that product's own market share and in which that market share appears as an additively separable linear term in the δ_{jt} index. These strong functional restrictions permit identification when other products' characteristics are valid instruments for a given product's market share.

Tastes for total quantities. Consumers may value products' total quantities—that is, their market shares times market sizes—instead of their market shares. This section considers identification in this case under the index restriction

$$\begin{aligned}\bar{\sigma}_{jt}(x_t, M_t s_t, \xi_t) &= \bar{\sigma}_j(\delta(x_t, M_t s_t, \xi_t)) \\ \delta_j(x_t, M_t s_t, \xi_t) &= x_{jt} + h_j(M_t s_t) + \xi_{jt}.\end{aligned} \tag{11}$$

I additionally impose the location normalization that, for each j, there is a known vector q_j such that $h_j(q_j) = 0$. When considering Assumption INVERT-MARKET in the context of the total quantities model, I use $\bar{\sigma}$ to denote the function of the index δ that appears on the right-hand side of the first equation in (11).

Although the total quantities model suffers from similar identification problems as the market shares model, cross-market variation in market size M_t allows for identification when this variation is assumed to be appropriately exogenous. Exogeneity of market size depends on mean-independence of consumer tastes from market size, which will fail to hold when consumers in differentially sized markets have different unobserved preferences. The following proposition characterizes the identification of the total quantities model.

Assumption NPIV-TOT (NPIV for total quantities model)**.** For each j, there is an observable random vector z_{jt} that satisfies following conditions:

(i) Exclusion restriction: $\mathbb{E}[\xi_{jt}|z_{jt}] = 0$ (almost surely).

(ii) Completeness condition: for all real-valued functions Γ such that $\mathbb{E}|\Gamma(s_t)| < \infty$, $\mathbb{E}[\Gamma(s_t)|z_{jt}] = 0$ (almost surely) implies $\Gamma(s_t) = 0$ (almost surely).

There are two sorts of available instruments in this setting: the BLP instruments—that is, characteristics x_{kt} of other products k other than j—and market size M_t. Combining these instruments yields z_{jt} of dimension J, which is the number of endogenous regressors in the nonparametric function that I will use NPIV-TOT to identify.

Last, I use several technical conditions to establish the identification of the total quantities model. These conditions restrict random elements' supports and impose the differentiability of structural functions to permit the use of calculus in identification analysis.

Assumption CALC-TOT (Technical conditions for total quantities model)**.** (i) For each j, h_j is differentiable.

(ii) $\bar{\sigma}$ is differentiable.

(iii) The support of $M_t s_t$ is convex.

Proposition 2. *Suppose that Assumptions INVERT-MARKET, NPIV-TOT, and CALC-TOT hold. Additionally suppose that at least one of the two following assumptions holds:*

(a) Own shares: $h_j(Ms)$ depends only on Ms_j for all j.

(b) Large support: for all s in the support of s_t, the support of M_t conditional on $s_t = s$ is $(0, \bar{M}(s)]$ for some $\bar{M}(s) > 0$.

Then, h is identified on the support of $M_t s_t$, $\bar{\sigma}$ is identified on the support of $\delta(x_t, s_t, \xi_t)$, and ξ_t is identified for all t.

Proof. See Appendix B.2. □

Proposition 2 establishes that the prospects for identification improve when choice probabilities depend on total quantities rather than market shares, although the proposition's conditions for identification are strong. The large support assumption is useful because it facilitates an identification-at-infinity argument relying on network externalities becoming irrelevant as the market size tends to zero. Appendix B.1 provides analysis of identification of the total quantities model in the case in which products' own quantities enter as additively separable linear terms in their δ indices. This functional restriction permits identification of the model.

3.3 Identification with microdata

Identification with microdata closely follows Berry and Haile (2024), with one key modification: market shares enter the individual choice probability function through network externalities. Identifying demand typically requires both instruments for prices, due to price endogeneity, and for market shares, to pin down substitution patterns. Absent network externalities, microdata eliminate the need for instruments for quantities because within-market variation identifies substitution patterns directly. When network externalities are present, however, shares become endogenous arguments of the choice probability function and must be instrumented.

The crucial distinction from the market-data setting is structural. With market data, shares enter the inverse demand equation in two distinct roles: mechanically, because inverse demand inverts shares into indices, and structurally, because network externalities depend on shares. These two roles cannot be separated empirically. Microdata isolate substitution patterns through within-market variation, so that only the structural network-effect channel remains to be instrumented. As a result, standard nonparametric instrumental variables techniques suffice to identify demand.

Throughout the analysis of identification with microdata, I impose the following index structure on consumer i's choice probability function:

$$\mathscr{s}_{ijt} = \sigma_j(\delta(w_{it}, \xi_t), x_t, s_t),$$

where $\mathscr{s}_{ijt}$ is consumer i's probability of choosing alternative j and δ is an index function that I specify as

$$\delta_j(w_{it}, \xi_t) = g_j(w_{it}) + \xi_{jt} \tag{12}$$

for $j \in \{1, \ldots, J\}$. The identification argument that follows does not make use of market characteristics x_t; therefore, I condition on these characteristics and suppress them in the notation. Additionally, the argument requires individual characteristics w_{it} of dimension J. Other individual characteristics may enter the model, but they do not play an important role in the identification analysis and are therefore omitted in my exposition.[4]

I now provide assumptions are adapted from Assumptions 2 and 3 in Berry and Haile (2024).

Assumption INV-DEMAND (Invertibility of demand)**.** The function $\sigma(\cdot, s_t)$ is injective on the support of $(w_{it}, \xi_t) \mid s_t$ almost surely.

Assumption INV-INDEX (Invertibility of index)**.** The function $g : \operatorname{supp} w_{it} \to \mathbb{R}^J$ is injective.

The model requires several normalizations to achieve identification. First, I impose $\mathbb{E}[\xi_t] = 0$ to rule out nonidentification due to the shifting of a constant between ξ_t and g. Second, I impose that $g(w_0) = 0$ for a known value $w_0 \in \operatorname{supp} w_{it}$. As a scale normalization, I impose $Dg(w_0) = I$, where D denotes the derivative operator. These last two normalizations are analogous to equations (7) and (8) in Berry and Haile (2024).

I additionally require the availability of variables z_t that satisfy the validity and relevance conditions required of instruments for market shares. These assumptions are the standard exclusion restriction and completeness condition of Newey and Powell (2003).

Assumption NPIV-EX (Exclusion restriction for NPIV)**.** $\mathbb{E}[\xi_t | z_t] = 0$ (almost surely).

Assumption NPIV-C (Completeness condition for NPIV)**.** For all real-valued functions Δ such that $\mathbb{E}|\Delta(s_t)| < \infty$, $\mathbb{E}[\Delta(s_t)|z_t] = 0$ (almost surely) implies $\Delta(s_t) = 0$ (almost surely).

Work on the identification of discrete-choice models often assumes that vectors of so-called *special regressors* are supported on the entirety of the relevant Euclidean space. Berry and Haile (2024) use an alternative and less restrictive assumption that serves a similar purpose in their identification proof. Assumption CPROB adapts this assumption to my setting. In this assumption, $\mathscr{s}$ denotes a consumer-specific choice probability whereas s denotes an aggregate market share. Additionally, Berry and Haile (2024) impose a nondegeneracy condition on the demand unobservables ξ_t that I adapt in Assumption NOND.

Assumption CPROB (Common choice probabilities)**.** There exists a choice probability vector $\mathscr{s}^*$ such that $\mathscr{s}^* \in \operatorname{supp} \mathscr{s}_{it} | \{s, \xi\}$, which is the support of $\mathscr{s}_{it}$ conditional on $s_t = s$ and $\xi_t = \xi$, for all $(s, \xi) \in \operatorname{supp}(s_t, \xi_t)$.

[4]See Berry and Haile (2024) for a treatment of additional individual characteristics, which they denote by Y_{it}.

Assumption NOND (Nondegeneracy)**.** There exists $s \in \operatorname{supp} s_t$ such that $\operatorname{supp} \xi_t \mid s_t = s$ contains an open subset of $\mathbb{R}^J$.

I provide some other technical conditions (TECH) and the proof of the following proposition in Appendix B.3.

Proposition 3. *Suppose that Assumptions INV-DEMAND, INV-INDEX, NPIV-EX, NPIV-C, CPROB, and NOND hold. Also suppose that the technical conditions TECH hold. Then, g is identified on the support of w_{it}, σ is identified on the support of $\{\delta(w_{it}, \xi_t), x_t, s_t\}$, and ξ_t is identified for all t.*

The proof parallels Berry and Haile (2024), except that shares now appear as endogenous arguments of the choice probability function and must be instrumented. Substitution patterns remain identified from within-market variation. A natural class of instruments exploits cross-market variation in the distribution of consumer characteristics. An example is $z_t = \mathbb{E}_t[w_{it}]$, the mean value of w_{it} in market t. In the context of the semi-parametric specification outlined by Section 2, another candidate instrument vector $z_t = [z_{1t}, \ldots, z_{Jt}]$ has components $z_{jt} = \Pr_t(w_{it}'\lambda_j + \varepsilon_{ijt} = \max_k w_{it}'\lambda_k + \varepsilon_{ikt})$ equal to the predicted shares of products j in market t when all market-level shifters (i.e., x_{jt}, s_t, and ξ_t) are removed from the utility equation. As long as F_t^w is mean-independent of ξ_t, these instruments will satisfy Assumption NPIV-EX. Unless $J_t = 1$, exogenous characteristics of other products are also available as instruments. See Section 3.4 for additional discussion of available instruments.

The argument used to identify the demand model in which market shares shift choice probabilities with microdata is straightforwardly adapted to the case in which total sales $M_t s_t$ shift choice probabilities. One needs only replace s_t with $M_t s_t$ where it appears in the choice probability function. In this case, M_t becomes available as an instrument as long as it satisfies Assumptions NPIV-EX and NPIV-C.

3.4 Relationship to empirical approaches

I now related my identification results to empirical practice. Timmins and Murdock (2007) draw on the methodological framework of Bayer and Timmins (2007) in estimating a discrete-choice model with network externalities using microdata and instrumental variables. In their framework, individual choice probabilities depend on market shares rather than total quantities. Their empirical setting consists of a single market, which precludes the use of instruments that exploit cross-market variation in consumer characteristics. My identification analysis formally justifies the use of instrumental variables in such microdata settings under appropriate exclusion and completeness conditions. It also clarifies how additional instruments could be obtained when multiple markets are observed, for example by exploiting cross-market variation in the distribution of consumer characteristics.

My analysis is also closely related to the identification strategy in Guiteras et al. (2025), who estimate demand for latrines in Bangladesh using a discrete-choice model with network externalities. Their data come from a randomized controlled trial that provided subsidies to a random subset of households, with variation in eligibility rates across neighbourhoods and villages. They use the share of households eligible for subsidies in a neighbourhood as an instrument for the share purchasing latrines. Such instruments are examples of instruments based on market-specific distributions of consumer characteristics, as highlighted in my identification analysis. My framework generalizes this insight by showing that cross-market variation in demographic characteristics that are not direct product attributes can provide identifying variation for network externalities under suitable exogeneity conditions.

4 Setting and data

I now illustrate the application of the article's identification insights to the estimation of a model of dating website choice. I estimate this model using data on the US dating websites industry in 2007–2008 from the Comscore Web Behavior Database. This dataset includes the browsing and online transactions records for a large panel of US households (91689 panelists in 2007 and 57817 panelists in 2008). Each panelist's activity is recorded for an entire year. Most individuals in the Comscore data appear in only one year. When an individual appears in multiple years, I treat that individual's records for each year as a separate panelist. Note that De Los Santos et al. (2012) find that individuals in Web Behavior Database are representative of online buyers in the US.

For each Comscore panelist, I observe each visited domain, the time of the visit, the duration of the visit, and the number of pages viewed. I also observe panelist characteristics including income group, educational attainment, race, age group, and ZIP code. I supplement the Comscore data with geographical and population data from the US Census Bureau.

Table 2 reports the most-used dating sites in 2007–2008. Each row of the "Share (%)" column reports the share of households in the Comscore data who spend at least five minutes on the indicated site. Most dating website usage in my sample is accounted for by a few major sites that appeal to broad audiences. These include `match.com`, `eharmony.com`, `pof.com` ("Plenty of Fish"), and `okcupid.com`, which are the sites that I focus on in my main analysis. Several smaller sites target various subpopulations (e.g. `catholicmatch.com` for Catholic users, `silversingles.com` for older users, and `farmersonly.com` for farmers). Whereas `eharmony.com` and `match.com` require payment for use, `pof.com` and `okcupid.com` are free to use. These free sites rely on advertising and paid premium features for revenue. The monthly prices of subscriptions for `eharmony.com` and `match.com` in 2007 were \$59.95 and \$34.99, respectively.

The markets in my analysis are geographical regions based on combined statistical areas (CSAs), which consist of counties that are economically and/or socially connected. I construct CSA/state pairs by dividing each CSA up into its parts that belong to different states. I then construct the geographical units underlying the markets in my analysis by adding each county that is within 50 miles of a CSA/state and that does not belong to a CSA/state to

Table 2: Most popular dating websites sites, 2007–2008

Site	Share (%)
match.com	8.61
eharmony.com	4.63
pof.com	1.93
chemistry.com	0.89
okcupid.com	0.82
matchmaker.com	0.54
lavalife.com	0.34
christianmingle.com	0.32
jdate.com	0.23
loveandseek.com	0.18
shaadi.com	0.16
badoo.com	0.16
zoosk.com	0.14
catholicmatch.com	0.07
farmersonly.com	0.04

Note: The "Share (%)" column provides the percentage of Comscore panelists spending ≥ 5 mins on the indicated site.

this county's closest CSA/state.

I assign each panelist in my estimation sample to either a primary site or the outside option. A panelist who visits at least five pages within a dating website, visits the site during at least two distinct sessions, and spends at least five minutes on the site qualifies as a user of that site. A panelist's primary site is the site on which the panelist spends the most time. To capture substitution into dating websites by users who did not previously use these websites, I specify an outside option consisting of panelists who visit a dating website but do not qualify as a user under the criteria above. I drop all other panelists from the sample used in my estimation procedure and counterfactual analysis. Furthermore, I drop markets with under 100 observed users.

Consumers are able to multi-home—i.e., use multiple websites—although few do so in practice. Among panelists who use at least one site, 81% use only one website. Online Appendix B provides additional information on the extent of multi-homing in the data. I rule out multi-homing in the empirical application.

Figure 3 provides evidence of cross-market variation in site market shares. In the presence of network externalities, a site's popularity in a region could be self-fulfilling in that a high number of users on a site explains the site's appeal to these users. In a market with the same exogenous characteristics, a site's lack of popularity could similarly be self-fulfilling. Thus, network externalities could lead a site to be popular in some regions but not others even in the absence of differences in site characteristics.

Figure 3: Cross-market variation in sites' market shares in 2007–2008

Notes: These plots display the 0.05, 0.10, 0.25, 0.50, 0.75, 0.90, and 0.95 quantiles of sites' market shares across markets in the 2007–2008 time period.

5 Estimation

Estimation of the model proceeds in two steps. In the first *microstep*, I estimate site/market-specific utility indices and preference loadings on consumer characteristics. In the second *market step*, I estimate the contribution of site characteristics and market shares to mean tastes for sites.

Microstep. The estimating equation of the microstep is

$$u_{ijt} = \delta_{jt}^{d(i)} + w_{ijt}'\lambda + \varepsilon_{ijt}, \tag{13}$$

where δ_{jt}^d are mean tastes for site j among members of demographic group d in market t and w_{ijt} are interactions of consumer characteristics with indicators for various sites j. Last, ε_{ijt} is assumed to follow a type 1 extreme value distribution. I estimate the parameters δ_{jt}^d and λ of (13) via maximum likelihood. Note that the microstep is the same for each of the share-type and quantity-type models.

Market step. The estimating equation of the market step of the share-type model is

$$\delta_{jt}^d = x_{jt}'\beta + f_j^d\left(s_t^{o}, s_t^1, \ldots, s_t^D; \gamma\right) + \xi_{jt}^d. \tag{14}$$

See Section 2 for a description of the terms appearing in this equation. The f_j^d network externality function is known up to the finite-dimensional parameter vector γ. Given that the utility indices δ_{jt}^d are not directly observed, I substitute estimates of these indices from the microstep for the true quantity in the estimation routine. The x_{jt} are firm-time period indicators whose coefficients I denote by ψ_j so that $x_{jt}'\beta = \psi_j$. Last, I estimate (14) using several different parametric forms of the network externality function $f_j^d(\cdot;\gamma)$, as I state explicitly in Section 6.

Following the identification analysis, I use Waldfogel instruments for the market shares in f_j^d. For each market share appearing in (14), I compute the predicted value of the market share based on the microstep estimates when (i) $\delta_{jt}^d = 0$ for all j, t, d and (ii) certain effects of individual characteristics w_{ijt} are also set to zero. I apply the same transformation to the predicted market shares when constructing the instruments as I do to the market shares when entering them into (14). To illustrate, when $D = 1$ and the specific estimating equation is

$$\delta_{jt} = x'_{jt}\beta + \gamma \log(s_{jt}) + \xi_{jt},$$

the Waldfogel instrument for s_{jt}

$$z_{jt} = \log\left(\frac{1}{M_t}\sum_{i=1}^{M_t} \frac{e^{\tilde{w}'_{ijt}\lambda}}{1 + \sum_k e^{\tilde{w}'_{ikt}\lambda}}\right),$$

where M_t is the sample size of consumers in market t and $\tilde{w}_{ijt}$ is a vector including a subset of the individual characteristics w_{ijt}. I similarly construct instruments for other specifications. As noted in my discussion of identification, Waldfogel instruments based on characteristics whose market-specific distributions directly affect consumer tastes violate the exclusion restriction required for these instruments to identify the model. Markets' distributions of characteristics including age, race, and education may directly affect taste for dating websites because people may directly value these characteristics in potential mates. The distribution of these traits in the population in a market may thereby shift people's desire to engage in dating, whether online or offline. With this potential failure of Waldfogel instruments in mind, I choose characteristics to include in $\tilde{w}_{ijt}$ whose market-specific distributions are least likely to directly affect consumers' tastes. These include an indicator for whether consumer i has broadband internet and the characteristics describing consumer i's internet usage.

In practice, I use estimates of the platform/market utility indices δ_{jt} from the microstep of my estimation procedure rather than the true values of these indices in estimating (14). Sampling error in estimation of δ_{jt} is asymptotically irrelevant, however, under the assumption that and the number of observations within each market t grows faster than the number of markets. I formally present this asymptotic argument in Online Appendix C.

The market step of estimating a quantity-type model proceeds similarly. For the quantity-type model, quantities replace market shares in (14) and predicted quantities are used in constructing instruments instead of predicted market shares.

5.1 Price sensitivity

To this point, I have ignored price competition. Estimating consumer price sensitivity is important for computing pricing equilibria in counterfactuals and for expressing welfare figures in dollar terms. But the fact that the dating websites in the sample charge uniform prices across geography means that I observe minimal price variation, which prevents me from estimating price sensitivity the alongside other preference parameters. Instead, I use my choice-model estimates and a model of price competition to estimate price sensitivity in an auxiliary estimation procedure. As previously mentioned, I use site/time indicators as the x_j in (14) and let ψ_j denote the fixed effect for site j. I then make the decomposition $\psi_j = \bar{\psi}_j - \alpha p_j$ and assume that the observed prices $\{p_j^*\}$ constitute a Nash equilibrium in pricing with marginal costs of zero in that

$$p_j^* = \arg\max_{p_j} \sum_t M_t \sigma_{jt}(p_j, p_{-j}^*; \alpha) p_j \qquad \forall j \text{ s.t. } p_j^* > 0. \tag{15}$$

The profit maximization problem in (15) gives rise to the first-order conditions (FOCs)

$$\sum_t \left[M_t \frac{\partial \sigma_{jt}}{\partial p_j}(p_j^*, p_{-j}^*; \alpha) p_j^* + \sigma_{jt}(p_j^*, p_{-j}^*; \alpha) \right] = 0 \tag{16}$$

which provided the basis of my estimation of α.

To compute the estimator $\hat{\alpha}$ of α, I substitute empirical analogues/estimates for population objects/parameters in each paid site's FOC (16) and then solve for α.[5]

Some sites are free to use. I do not include these sites' FOCs in the estimation of α and I assume that free sites remain free in my counterfactuals. Each paid site's FOC provides a separate estimate of α; my final estimator $\hat{\alpha}$ is the average of these site-specific estimates. I compute standard errors for $\hat{\alpha}$ using a parametric bootstrap that involves sampling from the estimated asymptotic distribution of the the parameters estimated in the market step of estimation. I include additional details on the estimation of α in Appendix C.[6]

6 Parameter estimates

This section reports and discusses my parameter estimates. I estimate the model using two different specifications of demographic groups: the baseline "Overall" specification wherein all consumer belong to the same demographic group and the alternative "Age" specification wherein consumers under the age of 35 belong to the first demographic group and all other consumers belong to the second demographic group.

I also estimate the model using several different specifications of the network externality function discussed in Section 5. The sites that consumers choose between are `eharmony.com`, `match.com`, `okcupid.com` and `pof.com`; choices to use other sites and a failure to use any site are grouped together in the outside option. The microstep of estimation involves a large number of parameters whose presentation I relegate to Online Appendix D. Many of the estimated parameters indicate significant taste differences across individuals with different

[5] These FOCs include price derivatives of market shares, which are not well defined without an assumption on how prices affect equilibrium selection in the presence of multiple equilibria. I assume that firms believe their market shares at counterfactual prices are given by the equilibrium surface around (χ_t, s_t) as defined in Section 2.1, where χ_t includes firms' prices in market t. I then use the price derivatives of this equilibrium surface as the price derivatives appearing in the FOCs underlying my estimation of α. The implicit function theorem provides an explicit form for these derivatives.

[6] To assess whether $\hat{\alpha}$ is a reasonable estimator, I consider its implications for a price response to monopolization. Under the counterfactual of Section 7.2 wherein `match.com` becomes a monopolist, it raises its price by 30.5% (see Table 8). This magnitude seems sensible.

observable characteristics.

Panel A of Table 3 displays the parameter estimates of an share-type model with the "Overall" demographic specification and the network externality function specification $f_j(s_t;\gamma) = \gamma \log(s_{jt})$. Instrumenting for market shares with the demographic instruments decreases the estimated coefficient of the network externality term relative to OLS. This reflects the mechanical positive correlation between the unobservables ξ_{jt} and market shares s_{jt}. Panel A also reports my estimate of α for this specification. The rows with names of websites (e.g. "eharmony") provide estimated site intercepts.

Panel B reports reports the first stage of the IV regression whose results are displayed in Panel A; in particular, it shows the results from a regression of $\widetilde{\log(s_{jt})}$ on $\tilde{z}_{jt}$, where $\tilde{\zeta}_{jt}$ denotes the residual of a regression of ζ_{jt} on the site-time indicators that are included as exogenous regressors in Panel A's IV regression. The first stage is strong with an F statistic of 8.3, indicating the relevance of the Waldfogel instruments.

We can use the results presented above to compute the value of an increase in a site's usership to its inframarginal users. The estimates for the baseline model of Table 3 imply that a 10% increase in the usership of a site is worth \$6.34 a month to an inframarginal user of that site, which was 18% of the site's price.

Tables 4 reports estimates from the "Age" demographic group specification under various specifications of the network externality function. In particular, column (1) of each table reports estimates from a specification in which consumers care only about the market share of a site within their own demographic group; column (2) reports estimates from a specification in which a consumer's tastes depend on the market share of a site within their own demographic group and the other demographic group; column (3) reports estimates from a specification in which a consumer's tastes depend on the market share of a site within their own demographic group only, but members of different demographic groups have different preferences for their own-group market shares; and column (4) reports estimates from a specification in which a consumer's tastes depend on the market share of a site both within their own demographic group and within the entire population. These tables suggest considerable homophily within groupings defined by age, as consumers more highly value market shares within their own age group than shares within the other group. Column (3) suggests that the strength of network externalities is similar within each demographic group—i.e., that there is not considerable heterogeneity in tastes for market shares between groups—although the own-group tastes for the older age group are not precisely estimated.

Online Appendix E provides estimates of a model in which overall quantities $M_t s_{jt}$ rather than market shares s_{jt} enter consumer utility. The results are similar to those for the baseline model.

Table 3: Market step parameter estimates – "Overall" demographic group specification

Panel A: Parameter estimates

	OLS	IV
$\log(s_{jt})$	0.99	0.68
	(0.02)	(0.15)
eharmony	0.13	-0.59
	(0.05)	(0.34)
match	-0.78	-1.28
	(0.04)	(0.24)
okcupid	-3.34	-4.71
	(0.10)	(0.65)
pof	0.48	-0.50
	(0.07)	(0.47)
p_j $(\hat{\alpha})$		0.0102
		(0.0041)

Panel B: First stage of IV regression

	$\widetilde{\log(s_{jt})}$
$\tilde{z}_{jt}$	0.94
	(0.32)
F	8.34

Table 4: Market step parameter estimates – "Age" demographic group specification

	(1)	(2)	(3)	(4)
Own-group $\log\left(s_{jt}^{d}\right)$	0.519	0.579	-	0.032
	(0.138)	(0.218)	-	(0.551)
Other-group $\log\left(s_{jt}^{d'}\right)$	-	-0.049	-	-
	-	(0.218)	-	-
Own-group $\log\left(s_{jt}^{\text{younger}}\right)$	-	-	0.607	-
	-	-	(0.189)	-
Own-group $\log\left(s_{jt}^{\text{older}}\right)$	-	-	0.406	-
	-	-	(0.181)	-
$\log\left(s_{jt}^{\text{overall}}\right)$	-	-	-	0.757
	-	-	-	(0.634)
p_{jt} $(\hat{\alpha})$	0.0027	0.0027	0.0031	0.0014
	(0.0007)	(0.0184)	(12.4134)	(0.0028)

7 Counterfactual analysis

7.1 Decomposition of variance in market shares

As illustrated by Figure 3, dating websites exhibit substantial variation in popularity across cities, despite minimal regional differences in site design. This raises the question of how websites establish market-specific competitive advantages. Variation in market shares could reflect differences in unobserved intrinsic preferences for particular platforms, captured by ξ_t, or cross-city differences in demographic composition. A third possibility is that network externalities generate local herding: once a platform gains an initial foothold in a city, its value increases endogenously with local participation, leading users to coordinate on different sites in different locations.

This mechanism parallels the insight of Glaeser et al. (1996) that social interactions can amplify small underlying differences across cities. In their context, higher crime rates encourage additional criminal activity, causing observed cross-city variation in crime to exceed what fundamentals alone would predict. Analogously, even modest initial differences in platform adoption may be magnified by network effects, producing large and persistent disparities in dating-site popularity across cities.

Table 5 reports a decomposition of cross-sectional variation in sites' market shares into components attributable to (i) unobserved taste heterogeneity ξ_{jt}, (ii) observed demographic heterogeneity, and (iii) network externalities. For each regime, I compute the standard deviation of market shares across geographic markets. To isolate cross-market variation rather than differences in average popularity across sites, I first demean each site's market share by subtracting its average share across markets.

In the "No net. ext." column, I eliminate the network externality component from consumers' indirect utilities and recompute equilibrium market shares in each market. Removing network externalities changes sites' mean utilities and therefore alters the market share of the outside option. To hold the outside-good share fixed, I add a constant $f^\dagger$ to each site's mean utility δ_{jt}, where $f^\dagger$ is chosen so that the implied market share of the outside good matches its observed value.

In the "No ξ_{jt}" column, I set all market-level unobserved mean taste shocks ξ_{jt} to zero. In the "No demo" column, I recompute market shares using the pooled distribution of demographic characteristics across all markets in the sample, rather than each market's own demographic distribution.

Table 5 reveals that network externalities explain most of cross-sectional variation in market shares. Indeed, removing market shares from consumers' indirect utilities reduces the cross-geography market share standard deviation by 58%. The table also establishes that unobserved differences in tastes across geography play a smaller but nonetheless significant role in explaining market share variation. Last, differences in tastes linked to observed demographic characteristics play a smaller but not insubstantial role.

Table 5: Decomposition of geographic variation in market shares

	No net. ext.	No ξ_{jt}	No demo.
Share of baseline SD	0.44	0.25	0.00
Δ	0.56	0.19	0.25

Table 6: Market shares in top markets under the baseline equilibrium

Market	`match.com`	`eharmony.com`
New York (NY)	0.156	0.134
Los Angeles (CA)	0.147	0.135
New York (NJ)	0.144	0.127
Miami (FL)	0.178	0.131
Chicago (IL)	0.141	0.158
Atlanta (GA)	0.136	0.182
San Jose (CA)	0.167	0.144
Dallas (TX)	0.145	0.137

7.2 Monopolization

Next, I use the estimated model to evaluate the effects of making `match.com` a monopolist. Throughout this exercise, I assume that prices constitute a Nash equilibrium between the paid firms. Given that each paid dating websites charged a single nation-wide price in the sample period, I constrain each site's price to be constant across markets.

The two effects on consumer welfare that I consider are *price response effects* and *network externality effects*. The former are defined as dollarized differences in expected utility between counterfactual equilibria with and without price responses. The latter are defined as differences in consumer enjoyment of network externalities between counterfactual equilibria with price responses and baseline equilibria. Overall differences in expected consumer utility between baseline and counterfactual equilibria also depend on changes in consumers' realized ε_{ijt} taste shocks. The monopolization counterfactual mechanically lowers expected utility by removing ε_{ijt} shocks for the removed sites in a way that depends critically on the assumed distribution of ε_{ijt}. To prevent these mechanical effects from driving the welfare results, and to highlight the roles of price responses and network externalities, I measure overall welfare effects using the sum of the price response effects and network externality effects rather than differences in expected utility.

Table 7 provides market shares in both the monopolization counterfactual and under the baseline market structure whereas Table 8 provides the change in `match.com`'s price. To summarize, `match.com` increases its prices in the absence of competition with other dating websites and nonetheless increases its market share. This increase in market share benefits `match.com`'s inframarginal users by increasing their utility from network externalities. Tab 9 provides mean welfare changes across markets. The results imply that although price increases from monopolization would reduce consumer welfare by \$5.95/consumer, this loss

Table 7: Market share changes in the monopoly counterfactual

	eharmony.com	match.com	okcupid.com	pof.com	Outside option
Baseline	0.150	0.147	0.012	0.040	0.651
Monopoly	0.000	0.191	0.000	0.000	0.809

Table 8: Price change for match.com in the monopoly counterfactual

Quantity	Value
Baseline	$44.66
Counterfactual	$58.28
Change	30.5%

would be significantly mitigated by gains from increased network coordination on match.com. Indeed, changes in the size of match.com's network boost welfare by $2.52/consumer, thus reducing the overall negative effect of monopolization by 42%. Put differently, ignoring the role of network externalities would lead the researcher to overstate the consumer losses from monopolization by 73%.

8 Conclusion

This paper analyzes the identification properties of a discrete-choice model with network externalities and uses such a model to study the market for dating websites. I use the model to assess the extent to which increased market concentration would benefit consumers who enjoy using the same platform as others. I find that network externalities are substantial and account for most variation in sites' market shares across geography in the US. Additionally, neither marginally increasing market concentration or monopolizing the dating website market boosts average consumer welfare: in both cases, any benefits from increased enjoyment of network externalities are more than offset by harms from price increases. Furthermore, increased enjoyment of network externalities from consolidation of the dating websites industry around a monopolist would significantly attenuate the consumer welfare harms from price increases associated with monopolization.

Table 9: Mean consumer welfare changes in the monopoly counterfactual ($/consumer)

Quantity	Value
Price effect	-5.95
Network externality effect	2.52
Combined effect	-3.43

Bibliography

Allende, Claudia. 2019. "Competition under social interactions and the design of education policies."

Angrist, Joshua. 2014. "The Perils of Peer Effects." *Labour Economics* 30 98–108.

Aradillas-López, Andrés. 2021. "Computing semiparametric efficiency bounds in discrete choice models with strategic-interactions and rational expectations." *Journal of Econometrics* 221 (1): 25–42.

Bayer, Patrick, Robert McMillan, and Kim Rueben. 2004. "An equilibrium model of sorting in an urban housing market." (10865): .

Bayer, Patrick, and Christopher Timmins. 2007. "Estimating equilibrium models of sorting." *The Economic Journal* 117 (518): 353–374.

Becker, Gary S., Kevin M. Murphy, and Harvard University Press eBook Package Backlist 2000-2013. 2000. *Social economics: market behavior in a social environment.* Cambridge, Mass: Belknap Press of Harvard University Press.

Berry, Steven, Amit Gandhi, and Philip Haile. 2013. "Connected Substitutes and Invertibility of Demand." *Econometrica* 81 (5): 2087–2111.

Berry, Steven, and Philip Haile. 2014. "Identification in Differentiated Products Markets Using Market Level Data." *Econometrica* 82 (5): 1749–1798.

Berry, Steven, and Philip Haile. 2016. "Identification in Differentiated Product Markets." *Annual Review of Economics* 8 27–52.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile prices in market equilibrium." *Econometrica* 63 (4): 841–890.

Berry, Steven T, and Philip A Haile. 2024. "Nonparametric identification of differentiated products demand using micro data." *Econometrica* 92 (4): 1135–1162.

Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin. 2009a. "Identification of peer effects through social networks." *Journal of Econometrics* 150 (1): 41–55.

Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin. 2009b. "Identification of peer effects through social networks." *Journal of Econometrics* 150 (1): 41–55.

Brock, William A., and Steven N. Durlauf. 2001a. "Discrete choice with social interactions." *Review of Economic Studies* 68 (2): 235–260.

Brock, William A., and Steven N. Durlauf. 2001b. "Interaction-based models." In *Handbook of Econometrics, Volume 5*, edited by Heckman, James, and Edward Leamer Volume 5. of Handbook of Econometrics 3297–3380, North-Holland.

Brock, William A., and Steven N. Durlauf. 2007. "Identification of binary choice models with social interactions." *Journal of Econometrics* 140 52–75.

Carrell, Scott E., Frederick V. Malmstrom, and James E. West. 2008. "Peer effects in academic cheating." *Journal of Human Resources* 43 (1): 173 – 207. 10.3368/jhr.43.1.173, Cited by: 128; All Open Access, Bronze Open Access.

De Los Santos, Babur, Ali Hortacsu, and Matthijs R. Wildenbeest. 2012. "Testing models of consumer search using data on web browsing and purchasing behavior." *American Economic Review* 102 (6): 2955–2980.

Durlauf, Steven, and Yannis Ioannides. 2010. "Social Interactions." *Annual Review of Economics* 2 451–478.

Farronato, Chiaria, Jessica Fong, and Andrey Fradkin. 2024. "Dog Eat Dog: Balancing Network Effects and Differentiation in a Digital Platform Merger." *Management Science* 70 (1): 464–483.

Fong, Jessica. 2024. "Effects of Market Size and Competition in Two-Sided Markets: Evidence from Online Dating." *Marketing Science* 43 (5): 971–985.

Glaeser, Edward L, Bruce I Sacerdote, and Jose A Scheinkman. 2003. "The social multiplier." *Journal of the European Economic Association* 1 (2-3): 345–353.

Glaeser, Edward L, Bruce Sacerdote, and Jose A Scheinkman. 1996. "Crime and social interactions." *The Quarterly journal of economics* 111 (2): 507–548.

Graham, Bryan S. 2008. "Identifying social interactions through conditional variance restrictions." *Econometrica* 76 643–660.

Graham, Bryan S. 2018. "Identifying and estimating neighborhood effects." *Journal of Economic Literature* 56 450–500.

Guiteras, Raymond, James Levinsohn, and Ahmed Mushfiq Mobarak. 2025. "Demand estimation with strategic complementarities: sanitation in Bangladesh." Unpublished working paper.

Hitsch, Günter J., Ali Hortacsu, and Dan Ariely. 2010. "Matching and sorting in online dating." *American Economic Review* 100 (1): 130–163.

Jullien, Bruno, Alessandro Pavan, and Marc Rysman. 2021. "Two-sided markets, pricing, and network effects." In *Handbook of Industrial Organization, Volume 4*, edited by Ho, Kate, Ali Hortacsu, and Alessandro Lizzeri Volume 4. of Handbook of Industrial Organization 485–592, Elsevier.

Manski, Charles. 1993. "Identification of endogenous social effects: the reflection problem." *Review of Economic Studies* 60 531–542.

Newey, Whitney, and James Powell. 2003. "Instrumental Variable Estimation of Nonparametric Models." *Econometrica* 71 (5): 1565–1578.

Rysman, Marc. 2004. "Competition between networks: a study of the market for Yellow Pages." *Review of Economic Studies* 71 (2): 483–512.

Smith, Lones. 2006. "The Marriage Model with Search Frictions." *Journal of Political Economy* 114 (6): 1124–1144.

Timmins, Christopher, and Jennifer Murdock. 2007. "A Revealed Preference Approach to the Measurement of Congestion in Travel Cost Models." *Journal of Environmental Economics and Management* 53 (2): 230–249.

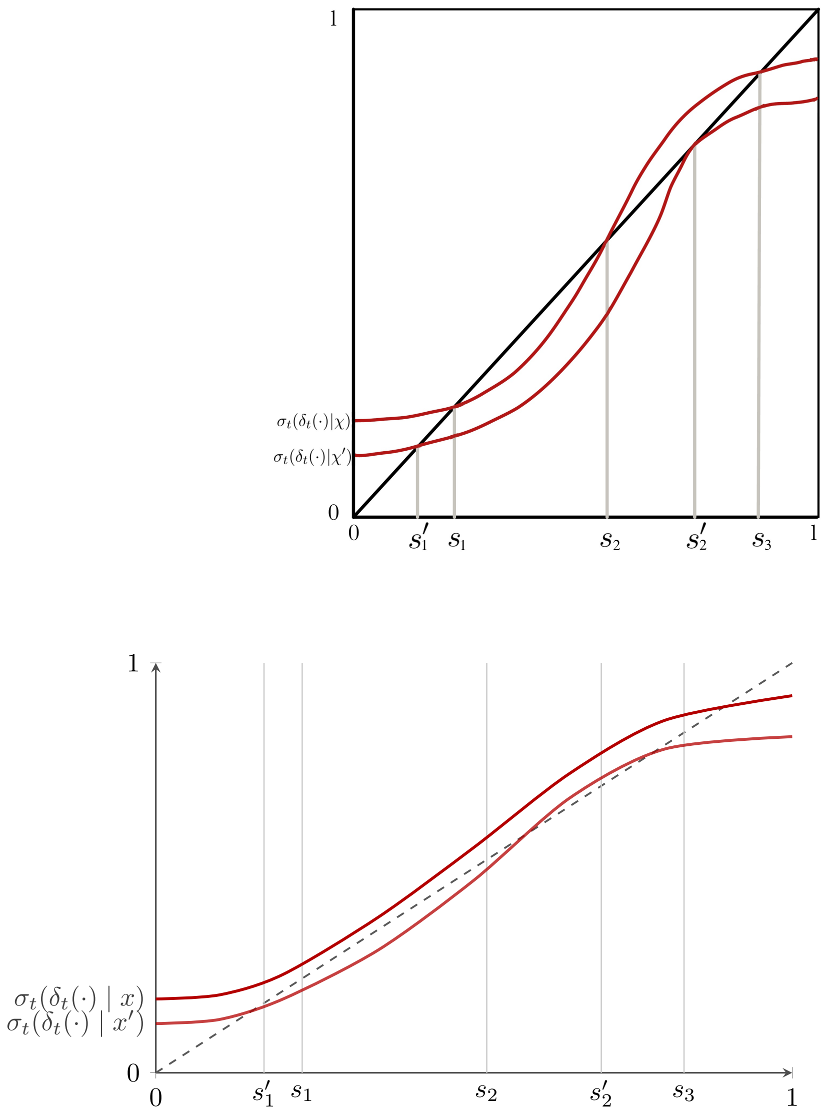
Waldfogel, Joel. 2003. "Preference Externalities: An Empirical Study of Who Benefits Whom in Differentiated-Product Markets." *The RAND Journal of Economics* 34 (3): 557–568.

Waldfogel, Joel. 2008. "The median voter and the median consumer: Local private goods and population composition." *Journal of Urban Economics* 63 (2): 567–582.

A Discussion of locally unique equilibria

This appendix discusses the concept of local uniqueness of equilibria as proposed in Section 2.1. Recall the condition that (4) must be nonsingular to ensure local uniqueness; I analyze this condition in the context of a binary logit model with positive network externalities. Figure 4 illustrates this model under two distinct values of the market characteristics χ_t. Equilibria in this model correspond to intersections between the s-shaped curve representing the mapping $s \mapsto \sigma_t(\delta_t(s)|\chi_t)$ and the 45-degree line. I interpret χ_t as the good's exogenous vertical quality; in the figure, decreasing χ_t from χ to χ' shifts the s-shaped curve downward until the equilibria s_2 and s_3 collapse into a single equilibrium s'_2. The matrix (4) is singular in this illustrative model when $\sigma_t(\delta_t(\cdot), \chi_t)$ is tangent to the 45-degree line, as happens when $\chi_t = \chi'$ at the equilibrium s'_2. We cannot define a unique equilibrium surface around (χ', s'_2) because there are two distinct equilibrium market shares nearby s'_2 when we marginally increase χ_t and no equilibrium market share near s'_2 when we marginally decrease χ_t. This illustration suggests that local uniqueness will fail only in knife-edge cases.

Figure 4: Multiple equilibria in a binary choice model with network externalities



B Identification appendix

B.1 Identification with market data using restrictions on network externalities

It is possible to establish identification with market data by significantly restricting the form of network externalities. Assumption SEPSHARE provides one way to restrict network externalities that helps with identification.

Assumption SEPSHARE (Separability of market share)**.** For each $j \in \mathcal{J}$, there is a function $\tilde{g}_j : \Delta^J \to \mathbb{R}$ such that

$$\delta_j(x_t, s_t, \xi_t) = \alpha s_{jt} + g_j(x_{jt}) + \xi_{jt},$$

where $\alpha \in \{-1, 1\}$ is known by the researcher.

Note that the restriction of α to the set $\{-1, 1\}$ is a scale normalization. The following assumption permits the application of a nonparametric instrumental variables argument to identify the model.

Assumption NPIV-MARKET (NPIV for market data)**.** Suppose $J > 1$. Let $z_{jt} = x_{-jt}$, i.e., a vector of all characteristics of products in market t excluding those of product j. These characteristics satisfy the following conditions for each $j \in \mathcal{J}$:

(i) Exclusion restriction: $\mathbb{E}[\xi_{jt}|z_{jt}] = 0$

(ii) Completeness condition: for all real-valued functions Γ such that $\mathbb{E}|\Gamma(s_t, x_{jt})| < \infty$, $\mathbb{E}[\Gamma(s_t, x_{jt})|z_{jt}] = 0$ (almost surely) implies $\Gamma(s_t, x_{jt}) = 0$ (almost surely).

Assumption NPIV-MARKET establishes characteristics of other products—that is, the BLP instruments—as valid instruments in my setting. Suppose that each x_{jt} has dimension d_x. This implies that the dimension of z_{jt} is $(J-1) \times d_x$, whereas the number of market shares is J. Therefore, $d_x > 1$ is required for the number of instruments to weakly exceed the number of market shares for which I seek instruments.

The following proposition characterizes identification under the assumptions above. The proposition's statement places an additional differentiability restriction on the inverse market share function $\bar{\sigma}^{-1}$; note that this inverse exists by Assumption INVERT-MARKET.

Proposition 4. *Suppose that Assumptions INVERT-MARKET, SEPSHARE, and NPIV-MARKET hold. Suppose additionally that the inverse market share function $\bar{\sigma}^{-1}$ is differentiable and that the support of s_t is path-connected. Then, g_j is identified on the full support of x_{jt} for each j, $\bar{\sigma}$ is supported on the full support of $\delta_j(x_t, s_t, \xi_t)$, and ξ_t is identified for each t.*

Proof. Inverting the equation

$$s_{jt} = \bar{\sigma}_j(s_{jt} + g_j(x_{jt}) + \xi_{jt}) \quad \forall j,$$

which is legal by Assumption INVERT-MARKET, yields

$$\bar{\sigma}_j^{-1}(s_t) = s_{jt} + g_j(x_{jt}) + \xi_{jt}$$

for each $j \in \mathcal{J}$. Here, I have assumed $\alpha = 1$, although the proof is essentially the same in the $\alpha = -1$ case. We can thus write

$$s_{jt} = \underbrace{\bar{\sigma}_j^{-1}(s_t) - g_j(x_{jt})}_{=:\kappa_j(s_t, x_{jt})} + \xi_{jt}. \tag{17}$$

Equation (17) falls within the NPIV framework: the left-hand side is observable and $\kappa_j(s, x_j)$ is an unknown function of (s, x_j). Consequently, the function κ_j is identified by the nonparametric instrumental variables argument of Newey and Powell (2003) under Assumption NPIV-MARKET. The identification of ξ_j follows immediately.

For the separate identification of $\bar{\sigma}_j^{-1}$ and $g_j(x_j)$, we require a location normalization; otherwise, we could shift a constant between these two functions without altering κ_j. As a location normalization, I impose that there is a $\bar{x}_j$ known to the researcher such that $g_j(\bar{x}_j) = 0$. Then, $\bar{\sigma}_j^{-1}(s) = \kappa_j(s, \bar{x}_j)$ and $g_j(x_j) = \kappa_j(s, x_j) - \bar{\sigma}_j^{-1}(s)$. Next, note that $\bar{\sigma}_j^{-1}$ is differentiable by assumption and that all partial derivatives of $\bar{\sigma}_j^{-1}$ are identified on the support of s_t by the separability of $\bar{\sigma}_j^{-1}$ and g_j in the definition of κ_j. Pick some $s^\dagger$ in the support of $s|\bar{x}_j$. Then, $\bar{\sigma}_j^{-1}(s^\dagger) = \kappa_j(s^\dagger, \bar{x}_j)$. Therefore, for any s in the support of s_t and a continuous curve r with a range contained in the support of s_t such that $r(0) = s^\dagger$ and $r(1) = s$ (which exists by virtue of the support being path-connected), the fundamental theorem of calculus for line integrals implies that

$$\begin{aligned}\int_0^1 \frac{\partial \bar{\sigma}_j^{-1}(r(\tau))}{\partial s} \cdot \nabla r(\tau) d\tau + \kappa_j(s^\dagger, \bar{x}_j) &= \bar{\sigma}_j^{-1}(s) - \bar{\sigma}_j^{-1}(s^\dagger) + \kappa_j(s^\dagger, \bar{x}_j) \\ &= \bar{\sigma}_j^{-1}(s),\end{aligned}$$

which identifies $\bar{\sigma}_j^{-1}$ on the entire support of s_t and consequently identifies $g_j(x_j) = \kappa_j(s, x_j) - \bar{\sigma}_j^{-1}(s)$ on the full support of x_j. The identification of $\bar{\sigma}$ on the support of $\delta(x_t, s_t, \xi_t)$ immediately follows from the identification of $\bar{\sigma}^{-1}$ on the support of s_t. □

The result above shows that a basic model of network externalities is identified by a substantive restriction on how market shares enter consumers' indirect utilities. This restriction is substantive in two ways: (i) it only allows a product's own market share to affect its indirect utility and (ii) it imposes a functional relationship between the product's own market share and its indirect utility, whereas this functional relationship may be what we seek to learn from the data.

The logic underlying the identification result above is readily applied to the variant of the model in which total quantities enter $\bar{\sigma}$ instead of market shares. The following proposition characterizes identification in this case.

Proposition 5. *Suppose that Assumptions INVERT-MARKET and SEPSHARE, and NPIV-MARKET hold, and that*

$$\delta_j(x_t, M_t s_t, \xi_t) = \alpha M_t s_{jt} + g_{jt}(x_{jt}) + \xi_{jt},$$

for all $j \in \mathcal{J}_t$ and for $\alpha \in \{-1, 1\}$, where α is known by the researcher. This is an adaptation of Assumption SEPSHARE to the total quantities model. Suppose additionally that the inverse market share function $\bar{\sigma}^{-1}$ is differentiable and that the support of s_t is path-connected. Then, g, $\bar{\sigma}$, and ξ_t are identified.

The proof is almost identical to that of Proposition 5. Note that, when M_t is assumed to satisfy the conditions required of z_{jt} by Assumption NPIV-MARKET, then M_t may be included in the z_{jt} vector; this reduces the requirement for BLP instruments. In the special case in which only a product's own market share appears in its utility index, there is only one endogenous regressor that requires an instrument. It is possible to deploy as many excluded instruments (one, M_t) as endogenous regressors in this case without using any BLP instruments.

B.2 Proof of Proposition 2

Proof. Assumption INVERT-MARKET implies the existence of an inverse $\bar{\sigma}^{-1}$ of the average choice probability function $\bar{\sigma}$ that satisfies, for each j,

$$\bar{\sigma}_j^{-1}(s_t) = x_{jt} + h_j(M_t s_t) + \xi_{jt}.$$

Re-arranging terms in the equation above yields a nonparametric regression equation:

$$x_{jt} = \kappa_j(s_t, M_t) - \xi_{jt}$$

for

$$\kappa_j(s, M) = \bar{\sigma}_j^{-1}(s) - h_j(Ms). \tag{18}$$

Assumption NPIV-TOT and the identification argument of Newey and Powell (2003) yields the identification of κ_j on the support of $\{s_t, M_t\}$. The two assumptions in the statement of the proposition provide two ways to separately identify $\bar{\sigma}^{-1}$ and h_j. Under assumption (a), we can write $h_j(Ms) = h_j(Ms_j)$ in a slight abuse of notation. In this case,

$$\frac{\partial \kappa_j}{\partial M}(s, M) = -h_j'(Ms_j)s_j \quad \Rightarrow \quad h_j'(Ms_j) = -\frac{1}{s_j}\frac{\partial \kappa_j}{\partial M}(s, M),$$

which shows that h_j' is identified on the support of $M_t s_{jt}$. Differentiation is legal in this context by Assumption CALC-TOT. The entire function h_j is identified under the location normalization provided in the main text that $h_j(q_j^*) = 0$ for a known number q_j^*; I take q_j^* as a number rather than a vector because h_j only depends on a scalar argument under assumption (a). Indeed, for q_t in the support of $M_t s_t$,

$$h_j(q_j) = \int_{q_j^*}^{q_j} h'(q)dq$$

by the fundamental theorem of calculus and $h(q_j^*) = 0$. Integration of h' over $[q_j^*, q_j]$ is justified by the convexity of the support of $M_t s_t$ as stipulated by Assumption CALC-TOT. The function $\bar{\sigma}_j^{-1}$ is then immediately identified on the support of s_t for each j, which implies the identification of $\bar{\sigma}$ on the support of $\delta(x_t, s_t, \xi_t)$. This completes the identification argument under assumption (a).

Now consider assumption (b). Note that

$$\frac{\partial \kappa_j}{\partial s}(s, M) = \nabla \bar{\sigma}_j^{-1}(s) - M \nabla h_j(Ms), \tag{19}$$

where ∇ is the gradient operator. Differentiation is legal in this context by Assumption CALC-TOT. We have

$$\lim_{M \downarrow 0} \frac{\partial \kappa_j}{\partial s}(s, M) = \nabla \bar{\sigma}_j^{-1}(s), \tag{20}$$

which shows how the large support assumption leads to identification of $\nabla \bar{\sigma}_j^{-1}$; the right-hand side of (20) is identified because (i) κ_j is identified on the full support of $\{s_t, M_t\}$ and (ii) the closure of this support includes $M_t = 0$ for all s_t. Note that the identification of ∇h_j consequently follows from (19). The levels of $\bar{\sigma}_j^{-1}$ and h_j are subsequently identified using the location normalization that $h_j(q_j^*) = 0$ for a known vector q_j^*. Indeed, the fundamental theorem of calculus for line integrals implies that for a differentiable path r with $r(0) = q_j^*$ and $r(1) = q_j$ in the support of $M_t s_t$,

$$\int_0^1 \nabla h_j(r(\tau)) \cdot r'(\tau) d\tau = h_j(q).$$

Integration of ∇h_j over a path r in the support of $M_t s_t$ is justified by the convexity of this support as stipulated by Assumption CALC-TOT. Given the definition of κ_j in (18), the identification of h_j on the support of $M_t s_t$ and the previously established identification of κ_j on the support of $\{s_t, M_t\}$ implies the identification of $\bar{\sigma}^{-1}$ on the support of s_t. The identification of $\bar{\sigma}^{-1}$ on the support of s_t straightforwardly implies the identification of $\bar{\sigma}$ on the support of $\delta(x_t, s_t, \xi_t)$. □

B.3 Proof of Proposition 3

Before presenting the proof, I provide technical conditions that adapt Assumption 5 of Berry and Haile (2024) to my setting.

Assumption TECH (Technical conditions)**.** The following conditions hold:

(i) $\operatorname{supp} w_{it}$ is open and connected;

(ii) g is uniformly continuous and continuously differentiable on $\operatorname{supp} w_{it}$;

(iii) $\sigma(\delta, s)$ is continuously differentiable with respect to δ for all $(\delta, s) \in \operatorname{supp}(\delta(w_{it}, \xi_t), s_t)$; and

(iv) $Dg(w)$ and $D_\delta \sigma(\delta, s)$ are nonsingular almost surely on $\operatorname{supp} w_{it}$ and $\operatorname{supp}(\delta(w_{it}, \xi_t), s_t)$, respectively.

Proof. The proof closely follows Berry and Haile (2024). Let $w^*(\mathscr{s}, s, \xi)$ be the vector of individual characteristics that give rise to choice probabilities $\mathscr{s}$ when market shares equal s and the vector of unobservable product qualities is equal to ξ. Such a vector exists for all $\mathscr{s}$

in the support of $\mathscr{s}$ conditional on $s_{it} = s$ and $\xi = \xi_t$.. Additionally, $w^*(\mathscr{s}, s, \xi)$ is unique by virtue of Assumptions INV-DEMAND and INV-INDEX. Next, let $\mathcal{W} = \operatorname{supp} w_{it}$ and let $\|\cdot\|$ be the Euclidean norm. By Lemma 1 in Berry and Haile (2024), there exists a $s \in \operatorname{supp} s_t$ and $\Delta > 0$ such that for all w and w' in $\operatorname{supp} w_{it}$ such that $\|w - w'\| < \Delta$, there exist a choice probability vector $\mathscr{s}$ and vectors ξ and ξ' in $\operatorname{supp} \xi_t \mid s$ such that $w = w^*(\mathscr{s}, s, \xi)$ and $w' = w^*(\mathscr{s}, s, \xi')$. This follows from Lemma 1 after substituting s_t in place of their price vector P_t and recalling that I condition on exogenous market characteristics x_t throughout my analysis. Next, Lemma 2 in Berry and Haile (2024) implies that there is a $\Delta > 0$ such that for almost all w and w' in $\operatorname{supp} w_{it}$ such that $\|w - w'\| < \Delta$, $(Dg(w))^{-1} Dg(w')$ is identified. By Lemma 3 in Berry and Haile (2024), g is identified.

Recall the index structure (presented here with the x_t characteristics suppressed in the notation): for $j \in \mathcal{J}_t$,

$$\begin{aligned} \mathscr{s}_{ijt} &= \sigma_j(\delta(w_{it}, \xi_t), s_t) \\ \delta_j(w_{it}, \xi_t) &= g_j(w_{it}) + \xi_{jt} \end{aligned} \tag{21}$$

The nonparametric regression equation used in identifying σ^{-1} is

$$g_j(w^*(\mathscr{s}^*, s_t, \xi_t)) = \sigma_j^{-1}(\mathscr{s}^*, s_t) - \xi_{jt} \tag{22}$$

We obtain this equation by σ with respect to its first argument in (21) at $w_{it} = w^*(\mathscr{s}, s_t, \xi_{it})$, where $\mathscr{s}^*$ is the common choice probability of Assumption CPROB; this inversion is justified by Assumption INV-DEMAND. The common choice probability $\mathscr{s}^*$ and the left-hand side of (22) are known. Therefore, (22) is a standard nonparametric regression equation with dependent variable $g_j(w^*(\mathscr{s}^*, s, \xi))$, nonparametric regression function $s \mapsto \sigma^{-1}(\mathscr{s}^*, s)$, and additive disturbance $-\xi_{jt}$. By the argument of Newey and Powell (2003), Assumptions NPIV-EX and NPIV-C identify each of the ξ_{jt} unobservables Given the choice probability function is

$$\mathscr{s}_{it} = \mathscr{s}(s_t, \xi_t, w_{it})$$

and both the left-hand side and each of the arguments of $\mathscr{s}$ is known, the function $\mathscr{s}$ is identified on its support. □

B.4 Identification of models with multiple demographic groups

I now consider the identification of models in which demographic-group-specific market shares appear in consumers' indirect utilities. I consider identification in two settings. In the first, which I call the submarket data setting, the researcher observes market shares and market sizes specific to each of the D demographic groups. In the second, which I call the microdata setting, the research observes microdata with individual choice probabilities and individual characteristics that vary within demographic groups.

I denote the market share of product j among consumers of demographic group d in market

t by s_{jt}^d and the measure of consumers belonging to demographic group d in market t by M_t^d. Let S_t be a $J \times D$ matrix whose dth column provides the market shares of the J inside goods among members of demographic group d in market t. I do not consider a model in which market shares rather than total quantities enter choice probability functions in the submarket data case because such a model suffers from the same identification problems as in the market data setting without multiple demographic groups. Instead, I begin by considering a total quantities model in the submarket data setting. Let $\bar{\sigma}_{j,d}$ denote the average choice probability function for product j and demographic group d; the equilibrium condition that determines market shares is

$$\bar{\sigma}_{j,d}(x_t, S_t M_t^\dagger, \xi_t) = s_{jt}^d, \tag{23}$$

which is the analogue of (3) for the setting with multiple market shares. In (23),

$$M_t^\dagger = \operatorname{diag} M_t = \begin{bmatrix} M_t^1 & 0 & \cdots & 0 \\ 0 & M_t^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & M_t^D \end{bmatrix}$$

so that

$$S_t M_t^\dagger = [s_t^1,\, s_t^2,\, \cdots,\, s_t^D] M_t^\dagger = [M_t^1 s_t^1,\, M_t^2 s_t^2,\, \cdots,\, M_t^D s_t^D].$$

I study identification under the following index structure:

$$\begin{aligned} \bar{\sigma}_{j,d}(x_t, S_t M_t^\dagger, \xi_t) &= \bar{\sigma}_j^d(\delta_j^d(x_t, S_t M_t^\dagger, \xi_t)) \\ \delta_j^d(x_{jt}, S_t M_t^\dagger, \xi_t) &= x_{jt}^d + h_j^d(S_t M_t^\dagger) + \xi_{jt}^d \end{aligned} \tag{24}$$

for each product j and each demographic group d. Here, $\bar{\sigma}_{j,d}$ is the average choice probability function specific to demographic group d.

The index structure (24) above restricts each x_{jt}^d to be scalar-valued, but the decomposition of x_{jt} into d-specific components is not necessarily an assumption; indeed, it remains possible to set $x_{jt}^d = x_{jt}$ for a single product characteristic x_{jt}. The assignment of a coefficient of one to x_{jt}^d is a scale normalization.

This model permits both the unobservables ξ_{jt}^d and the structural function g_j^d to vary across demographic groups. Under a generalization of Assumption INVERT-MARKET to the case of multiple demographic groups, we can invert d-specific market shares to obtain

$$\bar{\sigma}_{j,d}^{-1}(s_t^d) = x_{jt}^d + h_j^d(S_t M_t^\dagger) + \xi_{jt}^d.$$

Re-arranging terms yields a nonparametric regression equation:

$$x_{jt}^d = \underbrace{\bar{\sigma}_{j,d}^{-1}(s_t^d) - h_j^d(S_t M_t^\dagger)}_{=:\kappa_j^d(S_t, M_t)} + \xi_{jt}^d.$$

Note that κ_j^d includes $JD + D$ endogenous regressors. Candidate instruments include the D-dimensional M_t and the BLP instruments $\{x_{jt}^d : d \in D, j \in \mathcal{J}\}$, of which there are JD. Although we can reduce our instrument requirements by assuming that h_j^d does not depend on certain columns of S_t, i.e., by assuming that consumers do not care about demand within certain demographic groups, such an assumption also reduces the availability of available instruments. This is because the $x_{jt}^{d'}$ for excluded demographic groups d' do not shift s_t^d when group d consumers do not value $s_t^{d'}$.

Under an appropriate completeness condition and exclusion restriction, a nonparametric instrumental variables argument identifies κ_j^d and ξ_{jt}^d. We can then use, as in Proposition 2, additional conditions on h_j^d or the support of M_t to separately identify $\bar{\sigma}_{j,d}^{-1}$ and $\tilde{g}_j^d$.

I now discuss identification of in the microdata setting. Identification analysis in this setting is very similar to identification analysis in this setting without distinct demographic groups. I focus here on a model in which market shares enter the choice probability functions. Consumers in this model have characteristics $\{w_{ijt}\}_{j\in\mathcal{J}}$ that vary within demographic groups. I consider identification under the index structure

$$\begin{aligned} \mathscr{s}_{ijt} &= \sigma_{j,d(i)}(\delta(w_{it}, \xi_t), x_t, s_t) \\ \delta_j^d(w_{it}, \xi_t) &= g_j^d(w_{it}) + \xi_{jt}^d \end{aligned} \tag{25}$$

for $j \in \{1, \ldots, J\}$ and $d \in \{1, \ldots, D\}$, where $d(i)$ is consumer i's demographic group.

Under suitably generalized assumptions, Proposition 3 is generalized to identify the model with multiple demographic groups. Indeed, the g_j^d functions are identified by applying the argument of Berry and Haile (2024) for identifying g as summarized in the proof of Proposition 3 (see Appendix B.3) to each demographic group separately. With the g_d^j functions in hand, one can proceed with identification using nonparametric regression equations of the form (22) that are specific to individual demographic groups d. As noted in Appendix B.3, identifying the ξ_t unobservables in these equations is sufficient for identification.

C Price sensitivity estimation

This section provides additional details of the price-sensitivity estimation procedure described in Section 5.1. To begin, I define $\bar{\sigma}_{jt}(p, s)$ as the mean probability that a consumer in market t uses site j under prices p when she believes that the prevailing market shares are s. The mean in the definition of $\bar{\sigma}_{jt}$ is taken over t's distribution of consumer characteristics w_{ijt} and unobservables ε_{ijt}. The market shares s need not be the market shares consistent with the mean choice probabilities $\bar{\sigma}_j(p, s)$; they are just a member of $\Delta^J = \{s \in (0, 1)^J : \sum_{j=1}^J s_j \leq 1\}$. Let $\sigma_{jt}(p_t)$ denote the market shares that prevail under prices p_t. The function σ_t is implicitly defined by

$$\bar{\sigma}_t(p, \sigma_t(p)) = \sigma_t(p). \tag{26}$$

The implicit function theorem tells us that, under a nonsingularity condition,

$$D_p\sigma_t(p) = [I - D_s\bar{\sigma}_t(p, \sigma_t(p))]^{-1} D_p\bar{\sigma}_t(p, \sigma_t(p)). \tag{27}$$

When ε_{ijt} are iid type 1 extreme value random variables, the derivatives appearing on the right-hand side are straightforward to compute. Furthermore, in this case we can explicitly obtain an expression for α in terms of observables and estimated objects from the first-order condition (FOC) for a particular site j. This FOC is

$$\alpha = -\frac{\sum_t M_t\sigma_{jt}(p_t)}{\sum_t M_t(\tilde{D}_p\sigma_t(p_t))_{jj}p_j}.$$

for $\tilde{D}_p\bar{\sigma}_t = D_p\bar{\sigma}_t/\alpha$, which can be expressed solely in terms of market shares and parameters for which I obtain estimates in the two-step estimation of the consumer choice model. Let $\hat{d}_{jt}$ be the estimator of $\left(\tilde{D}_p\sigma_t(p_t)\right)_{jj}$ obtained by substituting estimates and empirical analogues of population objects into the form of $\tilde{D}_p\bar{\sigma}_t$ stated later in this appendix. My estimator of α is then

$$\hat{\alpha} = -\frac{1}{J}\sum_j \frac{\sum_t M_t s_{jt}}{\sum_t M_t\hat{d}_{jt}p_j}.$$

When the ε_{ijt} random variables are iid type 1 extreme value, the network externality function f_j depends only on s_j and is symmetric across j, and

$$\delta_j = \bar{\psi}_j - \alpha p_j + f(s_j) + \xi_j + \varepsilon_{ij},$$

we have

$$\begin{aligned}
\frac{\partial\bar{\sigma}_j}{\partial p_j} &= -\alpha\bar{\sigma}_j(1-\bar{\sigma}_j) \\
\frac{\partial\bar{\sigma}_j}{\partial p_k} &= \alpha\bar{\sigma}_j\bar{\sigma}_k \\
\frac{\partial\bar{\sigma}_j}{\partial s_j} &= \frac{\partial f}{\partial s_j}(s_j)\bar{\sigma}_j(1-\bar{\sigma}_j) \\
\frac{\partial\bar{\sigma}_j}{\partial s_k} &= -\frac{\partial f}{\partial s_k}(s_k)\bar{\sigma}_k\bar{\sigma}_j
\end{aligned}$$

Now note that

$$\tilde{D}_p\bar{\sigma}_t = \frac{1}{\alpha}D_p\bar{\sigma}_t$$

does not depend on α. This makes it convenient to write (27) as

$$D_p\sigma_t(p) = \alpha\underbrace{[I - D_s\bar{\sigma}_t(p, \sigma_t(p))]^{-1}\tilde{D}_p\bar{\sigma}_t(p, \sigma_t(p))}_{=:\tilde{D}_p\sigma_t(p)}.$$

Given market shares and the parameters of the consumer choice model, we can compute

$\tilde{D}_p\sigma_t(p)$ without knowledge of the price sensitivity α. We can write site j's FOC(16) as

$$\alpha \sum_t M_t(\tilde{D}_p\sigma_t(p_t))_{jj}p_j = -\sum_t M_t\sigma_{jt}(p_t).$$

when we assume the observed prices are equilibrium prices. Therefore,

$$\alpha = -\frac{\sum_t M_t\sigma_{jt}(p_t)}{\sum_t M_t(\tilde{D}_p\sigma_t(p_t))_{jj}p_j}.$$

Since the first-order condition holds for each j, we have

$$\alpha = -\frac{1}{J}\sum_j \frac{\sum_t M_t\sigma_{jt}(p_t)}{\sum_t M_t(\tilde{D}_p\sigma_t(p_t))_{jj}p_j}. \tag{28}$$

Let $\hat{d}_{jt}$ be the estimator of $\left(\tilde{D}_p\sigma_t(p_t)\right)_{jj}$ obtained by (i) substituting in observed market shares s_{jt} for $\bar{\sigma}_{jt}$ in the partial derivatives of σ_t with respect to market shares and (ii) substituting γ with an estimator $\hat{\gamma}$. Substituting in $\hat{d}_{jt}$ for $\tilde{D}_p\bar{\sigma}_t$ in (28) yields my estimator of α:

$$\hat{\alpha} = -\frac{1}{J}\sum_j \frac{\sum_t M_t s_{jt}}{\sum_t M_t\hat{d}_{jt}p_j}.$$

I now consider estimation of α under a more general share-type model with indirect utilities of the form

$$\delta_j^d = \bar{\psi}_j^d - \alpha p_j + f_j^d(s, s^1, \ldots, s^D) + \xi_j^d + \varepsilon_{ij},$$

Here, I allow network externalities to depend on the market shares of all demographic groups in addition to the market among particular demographic groups $d' \in \{1, \ldots, D\}$. I do not yet allow α to depend on d. Last, note that I explicitly allow the network externality term to depend both on overall market shares and demographic-specific market shares. The implicit function mapping prices into equilibrium market shares is given by the condition

$$\bar{\sigma}(p, \sigma(p)) = \sigma(p)$$

as before, but now σ includes a component for each site-demographic group pair. Site j's Bertrand-Nash equilibrium price is

$$p_j^* = \arg\max_{p_j} \sum_t \sum_d M_t^d \sigma_{jt}^d(p_j, p_{-j}^*; \alpha)p_j.$$

The corresponding FOC is

$$\sum_t \sum_d M_t^d \sigma_{jt}^d(p_j^*, p_{-j}^*; \alpha) + \sum_t \sum_d M_t^d \frac{\partial}{\partial p_j}\sigma_{jt}^d(p_j^*, p_{-j}^*; \alpha)p_j^* = 0.$$

The implicit function theorem provides a formula for $D_p\sigma_{jt}(p;\alpha)$:

$$D_p\sigma_t(p) = [I - D_s\bar{\sigma}_t(p,\sigma_t(p))]^{-1} D_p\bar{\sigma}_t(p,\sigma_t(p)).$$

Here, I is the $JD \times JD$ identity matrix. Also note that (i) the inverted matrix on the right-hand side is $JD \times JD$ whereas the price derivatives are $JD \times J$ and (ii) the specification above includes both share-type and quantity-type models since the f functions can implicitly depend on the populations of demographic groups within particular markets. I now provide the forms of the $D_s\bar{\sigma}$ and $D_p\bar{\sigma}$ functions under the multinomial logit parametric assumption:

$$\begin{aligned}
\frac{\partial\bar{\sigma}_j^d}{\partial p_j} &= -\alpha\bar{\sigma}_j^d(1-\bar{\sigma}_j^d) \\
\frac{\partial\bar{\sigma}_j^d}{\partial p_k} &= \alpha\bar{\sigma}_j^d\bar{\sigma}_k^d \\
\frac{\partial\bar{\sigma}_j^d}{\partial s_j^d} &= \bar{\sigma}_j^d(1-\bar{\sigma}_j^d)\left[\frac{\partial f_j^d}{\partial s_j}\frac{ds_j}{ds_j^d} + \frac{\partial f_j^d}{\partial s_j^d}\right] - \sum_{r\neq j}\bar{\sigma}_j^d\bar{\sigma}_r^d\left[\frac{\partial f_r^d}{\partial s_j}\frac{ds_j}{ds_j^d} + \frac{\partial f_r^d}{\partial s_r^d}\right] \\
\frac{\partial\bar{\sigma}_j^d}{\partial s_k^d} &= \bar{\sigma}_j^d(1-\bar{\sigma}_j^d)\left[\frac{\partial f_j^d}{\partial s_k}\frac{ds_k}{ds_k^d} + \frac{\partial f_j^d}{\partial s_k^d}\right] - \sum_{r\neq j}\bar{\sigma}_j^d\bar{\sigma}_r^d\left[\frac{\partial f_r^d}{\partial s_k}\frac{ds_k}{ds_k^d} + \frac{\partial f_r^d}{\partial s_k^d}\right] \\
\frac{\partial\bar{\sigma}_j^d}{\partial s_j^g} &= \bar{\sigma}_j^d(1-\bar{\sigma}_j^d)\left[\frac{\partial f_j^d}{\partial s_j}\frac{ds_j}{ds_j^g} + \frac{\partial f_j^d}{\partial s_j^g}\right] - \sum_{r\neq j}\bar{\sigma}_j^d\bar{\sigma}_r^d\left[\frac{\partial f_r^d}{\partial s_j}\frac{ds_j}{ds_j^g} + \frac{\partial f_r^d}{\partial s_r^g}\right] \\
\frac{\partial\bar{\sigma}_j^d}{\partial s_k^g} &= \bar{\sigma}_j^d(1-\bar{\sigma}_j^d)\left[\frac{\partial f_j^d}{\partial s_k}\frac{ds_k}{ds_k^g} + \frac{\partial f_j^d}{\partial s_k^g}\right] - \sum_{r\neq j}\bar{\sigma}_j^d\bar{\sigma}_r^d\left[\frac{\partial f_r^d}{\partial s_k}\frac{ds_k}{ds_k^g} + \frac{\partial f_r^d}{\partial s_k^g}\right].
\end{aligned}$$

Note that, since $s_j = \sum_d (M_t^d/M_t)s_j^d$, $ds_j/ds_j^d = M_t^d/M_t$. As in the simple model, the matrix $\tilde{D}_p\bar{\sigma}_t = \frac{1}{\alpha}D_p\bar{\sigma}_t$ does not depend on α and neither does

$$\tilde{D}_p\sigma_t(p) := \frac{1}{\alpha}\left[I - D_s\bar{\sigma}_t(p,\sigma_t(p))\right]^{-1} D_p\bar{\sigma}_t(p,\sigma_t(p)).$$

By construction, $D_p\sigma_t(p) = \alpha\tilde{D}_p\sigma_t(p)$. Let

$$\Delta_{jt}^d = \frac{1}{\alpha}\frac{\partial\sigma_j^d}{\partial p_j}(p^*;\alpha),$$

which does not depend on α by the analysis above. Therefore, we can write the FOC as

$$\alpha = -\frac{\sum_t\sum_d M_t^d\sigma_{jt}^d}{\sum_t\sum_d M_t^d\Delta_{jt}^d p_j^*}.$$

Substituting empirical analogues/estimates in for population objects provides an estimator $\hat{\alpha}$ of α.

C.1 Standard errors

Since the number of consumers in each market grows at a much faster rate than the number of markets, the first-order source of asymptotic variance in $\hat{\alpha}$ comes solely from asymptotic variance in our estimates from the market step of estimation. Thus, I compute standard errors for $\hat{\alpha}$ using a parametric bootstrap using the standard errors of my estimates from the market step of estimation.