Network externalities in the dating website industry*

Michael Sullivan

Department of Economics, Yale University

August 24, 2020

Abstract

Network externalities—i.e., effects on consumers' alternative-specific pay-offs of other consumers' decisions—arise in many economic settings. This paper considers the identification of discrete-choice models featuring network externalities and then uses such a model to analyze the role of network externalities in the market for dating websites in the United States. My identification analysis shows that discrete-choice models with network externalities are generally not identified with market-level data alone, but that microdata allows the researcher to construct instrumental variables that identify the model. Using rich online browsing data, I estimate a model of consumer choice of dating website and find evidence of considerable network externalities. In my preferred specification, an inframarginal user of a site values a 10% increase in the site's usership at \$11.11/month, which is about one third of the most popular site's price. I use my estimates to assess whether the observed extent of network externalities imply that increased market concentration would benefit consumers. I find that welfare losses from decreased variety and increased prices outweigh the gains from network externalities associated with a move from the observed market structure to monopoly.

*Email address: `m.r.sullivan@yale.edu`. I thank Steven Berry, Phil Haile, Katja Seim, and Mitsuru Igami as well as participants in Yale's IO Prospectus Workshop for for helpful feedback on this project.

1 Introduction

A decision-maker's pay-off from choosing a particular alternative is often affected by other decision-makers' choices; such effects are called network externalities. Examples of network externalities abound in economics. To provide a few, a driver may prefer uncongested roads, a web user may prefer messaging applications that are widely used in her social circle, and a consumer may prefer a smartphone that is fashionable due to its popularity in his region. Network externalities change standard analyses of competition and welfare in markets in which they affect consumer demand. As an example, consumers who enjoy using the same service as other consumers may prefer a more concentrated market despite harms from price increases and variety losses that this could entail. The possibility of gains from concentration is relevant in social media markets, where network externalities arise because consumers' enjoyment of a social media service depends on the selection of other users with whom they can interact on this service. Network externalities are central to the concept of a multi-sided platform given that they connect consumers on different sides of the platform.

In this paper, I study network externalities in the dating website industry. Some suggestive evidence for network externalities in this industry comes from the fact that sites' market shares exhibit considerable variation across geographically defined markets; this is consistent with consumers in different cities concentrating on different sites. But it is also possible that taste differences across markets induce cross-market variation in market shares. I estimate a model of consumer choice and price competition in this industry and use my estimates to explain the source of cross-market variation in sites' shares; these estimates suggest that network externalities are substantial and account for most this variation. Moreover, I use my estimates to quantify the value of network externalities to consumers, to provide evidence for age- and race-based homophily in dating website choice, and to assess whether consumers would prefer a monopoly in the dating website market (i.e. whether network externalities are strong enough to counteract the usual harms to consumers associated with market power). Under the estimates of my preferred specification, an inframarginal user of a site values a 10% increase in the site's usership at \$11.11/month, which is about one third of the most popular site's price. In addition, my estimates provide evidence for race- and age-based homophily in dating websites choice, i.e. that consumers especially value the presence of consumers of their same racial group and age group on their selected dating website. I also use my estimates to assess whether the observed extent of network externalities imply that increased market concentration would benefit consumers. Counterfactual analysis conducted with my estimated model suggests that welfare losses from decreased variety and increased prices outstrip the gains from network externalities associated with a move to monopoly in the dating website industry.

One of this paper's primary contributions is to characterize the identification properties of a class of discrete-choice models with network externalities. Network externalities complicate the identification of discrete-choice models because they typically amplify or dampen the effects of choices' characteristics on consumers' choice probabilities. To illustrate, a dating website's price decrease may increase its usage both because (i) consumers hate paying for dating websites and (ii) consumers value the additional users attracted to the site by the price decrease. Disentangling these two effects is the primary identification problem that this paper considers. I formally show that, unlike standard discrete-choice models, a general discrete-choice model in which market shares enter consumers' alternative-specific payoffs is not identified with market-level data alone. The most promising way to identify the model with network externalities involves using microdata, i.e. data on individual consumers' characteristics and decisions. The primary identification insight here is that, with market data alone, it is impossible to vary product characteristics while holding market shares fixed to identify their effects on consumers' choice probabilities. This is the case because market shares and choice probabilities are one-and-the-same in the market-data setting. With microdata, we can use within-market variation in consumer characteristics to identify the effects of these characteristics on choice probabilities. Varying a market's distribution of consumer characteristics has two effects on market shares: a direct effect reflecting that a consumer's tastes depend on that consumer's characteristics and an indirect effect reflecting that the direct effect changes market shares and consumers' choices depend on these shares via network externalities. Within-market variation identifies the direct effect and cross-market variation identifies the overall effect; combined, these sources of variation identify the indirect effect of network externalities.

Another contribution of this paper is to use a choice model featuring network externalities alongside a model of supply-side competition to study a social media industry (i.e. the dating websites industry) via counterfactual analysis. Although increasing market concentration often hurts consumers by increasing prices and decreasing product variety or quality, it can benefit consumers who like belonging to the same platform as others. Additionally, the entry of sites appealing to niche audiences can benefit members of these audiences while making inframarginal users of the primary existing sites worse off by reducing their market shares. In both examples, the overall welfare effect of the change in the market environment is ambiguous. I use the estimated model to counduct various counterfactual analyses intended to reveal the welfare effects of increased concentration and of the entry of niche sites. The analyses considered here may be useful in understanding other online industries with network externalities, e.g. business directories or job-listing websites.

1.1 Relation to the literature

Several papers consider the use of discrete-choice models in empirically detecting network externalities. Bayer and Timmins (2007) propose such a model and a procedure for estimating its parameters. Timmins and Murdock (2007) use Bayer and Timmins's model and estimator to study the site choice of recreational fishers in Wisconsin. These papers consider an econometric endogeneity problem arising from the fact that, in the presence of network externalities, consumers' choices depend on both market shares and demand unobservables that shift market shares. I consider the same problem in this paper and relate my own identification analysis to the instrumental variables (IV) approach of Bayer and Timmins and Timmins and Murdock. Guiteras et al. (2019) use a similar discrete-choice model with network externalities to study demand for latrines in Bangladeshi villages. Like Bayer and Timmins and Timmins and Murdock, Guiteras et al. use IVs in their estimation procedure. I also draw on their IV approach in my identification analysis and relate my analysis to their instruments.

My paper also relates Bayer et al. (2004), who study residential choice with sorting induced by the fact that households care about the characteristics of their neighbours. Bayer et al. (2004) note that models with network externalities generally support multiple equilibria; the authors avoid the problem of selecting equilibria in their counterfactual experiments by emphasizing that the equilibria of their model are locally unique in a neighbourhood of

Table 1: Number of panelists in Comscore Web Behavior Database

	2007	2008	2017	2018
N. panelists	91689	57817	93808	83666

the observed market characteristics and considering outcomes under counterfactual market characteristics that are nearby the observed market characteristics. I use a similar approach.

Two empirical studies in industrial organization that relate to my work are Rysman (2004) and Allende (2019). Rysman (2004) studies of network externalities in the market for Yellow Pages directories. In Rysman's model, consumers like adverts in Yellow Pages directories, which implies that an advertiser's decision to place ads in a directory benefits other advertisers in that directory by increasing consumer usage of the directory. Rysman asks whether these network externalities imply that reduced competition among publishers (e.g. moving from the observed extent of competition to monopoly) could be welfare enhancing. I address an analogous question for the dating website industry in this paper. My approach differs from Rysman's in several ways. First, I focus on only the consumer side of the dating website market and not on advertisements so that, for my purposes, dating websites are one-sided platforms. Second, I focus on consumer welfare rather than the welfare of advertisers and publishers. Last, Allende (2019) studies competition among schools in Peru using a model in which households value the socio-economic status (SES) composition of students at schools. Her demand model and equilibrium definition is similar to that developed in my paper.

2 Setting and data

In this paper, I study the dating website industry in the United States (US) between 2007 and 2018. My primary source of data is the Comscore Web Behavior Database (Comscore 2007, 2008, 2017, 2018). This dataset includes the browsing records and online transactions for a large panel of US households. Comscore records these data using a proxy server through which all of its panelists' online activity is routed. Each panelist's activity is recorded for an entire year of participation. Table 1 reports the number of Comscore panelists in each year of my data. Most Comscore panelists appear in only one year of the data. When a panelist appears in multiple years of data, I treat the panelist as a separate consumer for each year of data in which she appears. Note that De Los Santos (2012) find that individuals in Web Behavior Database are representative of online buyers in the US.

For each household in the Comscore data, I observe each domain visited by the consumer, the time of the visit, the duration of the visit, and the number of pages viewed by the consumer during the visit. I also observe various consumer 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 both the 2007–2008 and 2017–2018 parts of the Comscore data. 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"), `okcupid.com`, and

Table 2: Most popular dating websites sites

2007–2008		2017–2018	
Site	Share (%)	Site	Share (%)
match.com	8.61	match.com	1.41
eharmony.com	4.63	pof.com	1.37
pof.com	1.93	zoosk.com	0.55
chemistry.com	0.89	okcupid.com	0.37
okcupid.com	0.82	eharmony.com	0.27
matchmaker.com	0.54	tinder.com	0.20
lavalife.com	0.34	badoo.com	0.16
christianmingle.com	0.32	elitesingles.com	0.09
jdate.com	0.23	chemistry.com	0.07
loveandseek.com	0.18	twoo.com	0.06
shaadi.com	0.16	farmersonly.com	0.06
badoo.com	0.16	christianmingle.com	0.06
zoosk.com	0.14	loveandseek.com	0.04
catholicmatch.com	0.07	shaadi.com	0.04
farmersonly.com	0.04	silversingles.com	0.03

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

Table 3: Prices of monthly subscription at paid dating websites in 2007

Site	Price
`eharmony.com`	\$59.95
`match.com`	\$34.99

`zoosk.com`, which are the sites that I focus in my main analysis. A smaller but not insubstantial amount of usage is accounted for by sites appealing to subpopulations defined by religious affiliation (e.g. `catholicmatch.com` for Catholic users), age (e.g. `silversingles.com` for older consumers), or occupation (e.g. `farmersonly.com` for farmers).

Of the sites on which I focus in my empirical analysis, `eharmony.com` and `match.com` require payment for use whereas `pof.com`, `okcupid.com`, and `zoosk.com` are free to use. These free sites rely on advertising and paid premium features for revenue. Table 3 provides the monthly prices of subscriptions for each of `eharmony.com` and `match.com` in 2007.

The markets in my analysis are geographical units during particular time periods. The geographical units are 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. The New York City CSA, for example, has parts in New York, Connecticut, New Jersey, and Pennsylvania. The CSA/states that I form from this CSA are New York (NY), New York (CT), New York (NJ), and New York (PA). 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 but that does not belong to a CSA/state to its closest CSA/state.

I assign each Comscore panelist to either a primary site or the outside option. A panelist who visits a dating website during at least two distinct sessions qualifies as a user of that site. A panelist's primary site is the site on which the panelist spends the most time among the sites of which the panelist is a user. All panelists who do not use any site are taken as choosing the outside option. After determining whether each panelist is a dating website user, I drop all markets from my analysis with under 200 observed users.

2.1 Relationship between local population and dating website usage

If consumers prefer to use dating websites that are popular in their localities, then we may expect dating websites to be especially popular in high-population localities with higher potential numbers of dating website users. The correlation between a measure of the local population around a consumer's residence and the consumer's dating website usage may also be affected by the fact that more populous areas could have more offline dating opportunities.

To check whether patterns in the Comscore data are suggestive of either of the two aforementioned phenomena, I regress measures of dating website usage on a measure of local population. I define a consumer's local population as the combined population of ZIP codes whose geographic centres are within 10 miles of the geographic centre of the consumer's own ZIP code. Table 4 reports some of the quantiles of this variable in my sample of Comscore panelists in 2007–2008. Table 5 reports the results of the usage regressions whereas Figure 1 plots the coefficients and 95% confidence intervals from Panel A of this table. The results in Panel A correspond to regressions where indicators for population ranges are the only regressors whereas the results in Panel B correspond to regressions in which I also control for income, race, educational attainment, internet speed, age, household size, and census region. The row, e.g., "Pop.: 0.10q to 0.25q" refers to an indicator for a user's local population falling between the 0.10 and 0.25 quantiles of this variable. The first row of each panel provides the dependent variable in the regression. The "usage indicator" variable takes on a value of one if the household viewed at least 10 pages on a dating website across at least 5 browsing sessions and spent at least 5 minutes on dating websites. Otherwise, it takes on a value of zero. Here, the dating websites included in the analysis are `eharmony.com`, `match.com`, `pof.com`, and `okcupid.com`. "Duration" is the total time in minutes that the user spent on these dating websites. "Pages viewed" is the number of pages on these websites that the user viewed. "Sessions" is the number of distinct browsing sessions in which the user visited one of these dating websites. For all regressions except the "usage indicator" regressions, I dropped users with durations, number of pages viewed, or number of sessions above these variables' respective 0.99 quantiles.

The results in Table 5 are not estimates of the effect of local population on dating website usage given that people who live in high population areas are unobservably different than people who live in less populous areas and these differences could plausibly effect dating website usage. With this being said, the result that dating website usage seems to initially rise in local population before falling is compatible with both (i) network externalities and (ii) offline dating opportunities correlated with local population competing with online dating.

Last, Figure 2 provides evidence of cross-market variation in sites' market shares. If network externalities exist in the market for dating websites, then a site's popularity in a region could be self-fulfilling in the sense that the high number of users on the site makes the site attractive to these users. In a market with the same exogenous characteristics, a site's

Table 4: Distribution of local population measure

τ	τ^{th} quantile
0.10	20024
0.25	66247
0.50	264507
0.75	670983
0.90	1291597

unpopularity could be self-fulfilling: a site with few users is unattractive to potential users when these users value the site for the possible connections they can make on the site. Thus, network externalities could lead a site to be popular in some regions but not others even without differences in site characteristics or consumer tastes across regions. Later in this paper, I will decompose the variation in sites' market shares across geographically defined markets into a part reflecting network externalities and parts reflecting taste heterogeneity across regions.

3 Model

This section proposes a discrete-choice model of demand featuring network externalities. Although I consider a semi-nonparametric generalization of the model in Section 4's discussion of identification, the remainder of the current section describes the parametric models that I ultimately bring to my data.

3.1 Simple model

Before outlining my model in full detail, I present a simple version of the model to emphasize its principal departure from standard discrete-choice models of demand. Consider a market in which each consumer i faces a choice between J discrete alternatives. In a standard discrete-choice model of demand, the consumer chooses the product j that maximizes u_{ij}, where

$$u_{ij} = x_j'\beta + \xi_j + \varepsilon_{ij},$$

where x_j are characteristics of product j, ξ_j is the unobserable quality of product j, and ε_{ij} is the idiosyncratic part of consumer i's taste for good j (see, e.g., Berry 1994). We include network externalities in the model by writing

$$u_{ij} = x_j'\beta + \gamma s_j + \xi_j + \varepsilon_{ij},$$

where s_j is the market share of good j, i.e. the probability of choosing good j integrated over all consumers i in the market. If $\gamma > 0$, then consumers enjoy choosing the same product as other consumers, i.e. there is agglomeration. If $\gamma < 0$, consumers dislike when others choose the same product, i.e. there is congestion.

Table 5: Local population and dating website usage

Panel A: Basline results

	Usage indicator	Duration	Pages viewed	Sessions
Pop.: under 0.10q	0.0590	8.78	10.67	1.21
	(0.0020)	(0.47)	(0.68)	(0.04)
Pop.: 0.10q to 0.25q	0.0656	10.43	13.63	1.36
	(0.0017)	(0.38)	(0.55)	(0.03)
Pop.: 0.25q to 0.50q	0.0671	10.61	14.43	1.36
	(0.0013)	(0.29)	(0.43)	(0.03)
Pop.: 0.50q to 0.75q	0.0686	10.88	15.54	1.36
	(0.0013)	(0.29)	(0.43)	(0.03)
Pop.: 0.75q to 0.90q	0.0677	10.18	14.46	1.35
	(0.0017)	(0.38)	(0.55)	(0.03)
Pop.: over 0.90q	0.0626	9.35	12.80	1.28
	(0.0020)	(0.47)	(0.68)	(0.04)
R^2	0.066	0.033	0.029	0.065
N	147092	145245	145245	145245

Panel B: Results with controls

	Usage indicator	Duration	Pages viewed	Sessions
Pop.: under 0.10q	0.0690	12.20	11.99	1.28
	(0.0043)	(0.97)	(1.42)	(0.09)
Pop.: 0.10q to 0.25q	0.0743	13.88	14.65	1.40
	(0.0041)	(0.94)	(1.37)	(0.08)
Pop.: 0.25q to 0.50q	0.0749	14.15	15.23	1.37
	(0.0041)	(0.92)	(1.34)	(0.08)
Pop.: 0.50q to 0.75q	0.0764	14.54	16.37	1.36
	(0.0041)	(0.93)	(1.36)	(0.08)
Pop.: 0.75q to 0.90q	0.0764	13.95	15.53	1.37
	(0.0043)	(0.97)	(1.41)	(0.09)
Pop.: over 0.90q	0.0700	13.03	13.90	1.28
	(0.0043)	(0.97)	(1.42)	(0.09)
R^2	0.071	0.034	0.031	0.07
N	147092	145245	145245	145245

Figure 1: Local population and dating website usage

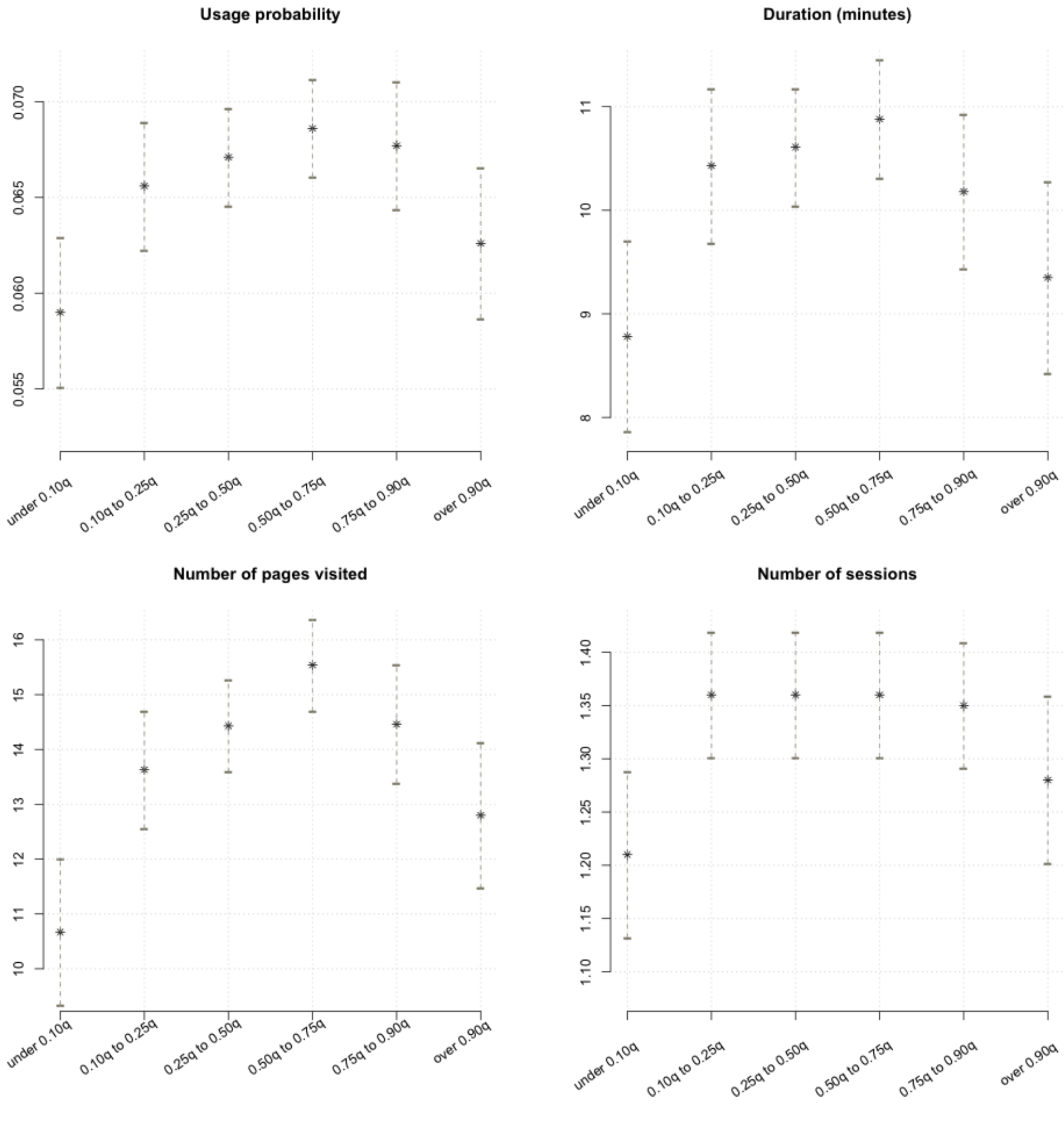
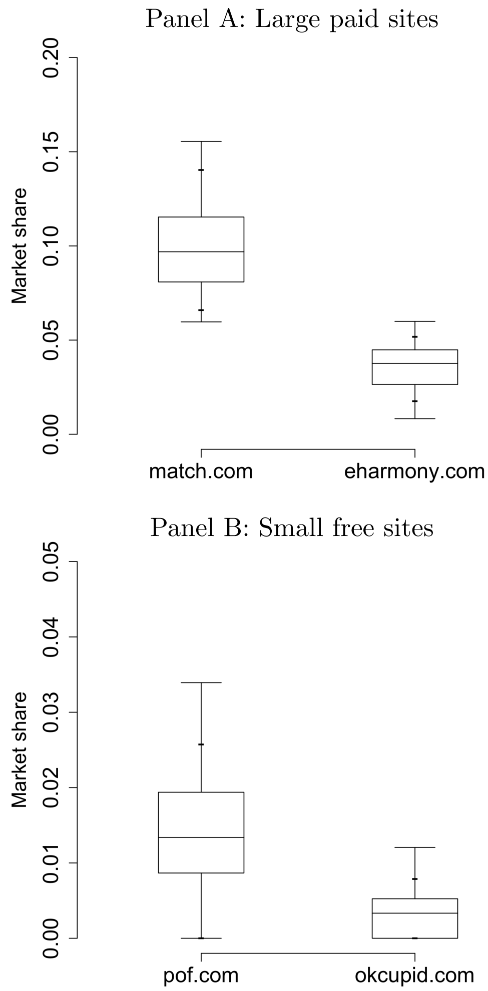


Figure 2: Cross-market variation in sites' market shares in 2007–2008



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.

3.2 Full model

I now describe my model in full detail. This model features markets t, each with a measure M_t continuum of consumers. Each consumer chooses between J_t products and an outside option. Let $\mathcal{J}_t$ denote the set of products *excluding* the outside option. Each consumer belongs to one and only one of D demographic 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 race (e.g. white and non-white) or age (e.g. under and over 35 years old) to allow preferences for other users to depend on the race or age of those users. Consumer i selects the product j that yields the greatest indirect utility u_{ijt}, where

$$\begin{aligned} u_{ijt} &= x_{jt}'\beta + w_{ijt}'\lambda + f_j^{d(i)}\left(s_t^o, s_t^1, \ldots, s_t^D\right) + \xi_{jt}^{d(i)} + \varepsilon_{ijt} \\ u_{i0t} &= h(M_t) + \varepsilon_{i0t}. \end{aligned} \tag{1}$$

In the expressions above, $s_t^d = [s_{1t}^d, \ldots, s_{J_t t}^d]'$ are market shares among consumers of group d; $s_t^o = \sum_d (M_t^d/M_t) s_t^d$ are the products' overall market shares; x_{jt} are observable characteristics of product j in market t; w_{ijt} are observable characteristics of consumer i in market t that may also depend on the product j; ξ_{jt}^d is the unobservable quality of product j in market t for demographic group d; ε_{ijt} is the unobservable aspect of consumer i's idiosyncratic tastes for j. Let F_t^w denote the distribution of $\{w_{ijt}\}_{j\in\mathcal{J}_t}$ in market t. Additionally, f_j^d is a function unknown to the econometrician (or known up to a finite-dimensional parameter). Last, the $h(M_t)$ term allows the value of the outside option to depend on the market size (in the dating website case, offline dating opportunities may depend on market size). For the remainder of the paper, I 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 $d \in \{1, \ldots, D\}$.)

In a discrete-choice model of demand without network externalities, we obtain market shares simply by taking the average of consumers' choice probabilities. When we introduce network externalities, market shares are instead fixed points of a "market share function." For simplicity of exposition, I now focus on the case of $D = 1$ so that there is a single market share for each product; the generalization to the case of $D > 1$ is straightforward and I consider specifications with $D > 1$ in my empirical analysis. The exogenous objects characterizing a market in this model are the distribution of demographic chararcteristics F_t^w, markets' distributions of the ε_{ijt} random variables, the observed product 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 solutions of the equations

$$s_{jt} = \Pr_t\left(j = \arg\max_{k\in\mathcal{J}_t} u_{ikt} \mid \chi_t\right) \qquad \forall j \in \mathcal{J}_t. \tag{2}$$

Note that the $\Pr_t$ operator integrates over both the ε_{ijt} deviates and the individual characteristics w_{ijt}; the operator is indexed by t because the distribution of w_{ijt} may vary across markets. Throughout this paper, I assume that consumers' w_{it} and ε_{ijt} values are drawn independently of each other and independently of the any mechanism determining the selection of equilibrium market shares (there are generally multiple equilibrium market shares in the model; see Subsection 3.4).

I now define the aforementioned *market share function* $\sigma_t(\cdot|\bar{\chi}) : \mathbb{R}^J \to \Delta^J$, where $\Delta^J \coloneqq \{s \in (0,1)^J : \sum_{j=1}^J s_j \leq 1\}$, as

$$\sigma_{jt}(\delta|\bar{\chi}) = \Pr_t\left(j = \arg\max_{k\in\mathcal{J}} \delta_k + w'_{ikt}\lambda + \varepsilon_{ikt} \mid \chi_t = \bar{\chi}\right).$$

This definition allows us to write (2) as

$$s_{jt} = \sigma_{jt}(\delta_t(s_t)|\chi_t) \qquad \forall j \in \mathcal{J}_t \tag{3}$$

where

$$\delta_{jt}(s_t) = x'_{jt}\beta + f_j(s_t) + \xi_{jt}. \tag{4}$$

I include s_t in $\delta_{jt}(s_t)$ to emphasize that equilibrium market shares s_t are defined as fixed-points of the mapping $s_t \mapsto \sigma_t(\delta_t(s_t)|\chi_t)$.

Although the ξ_{jt} unobservables are exogenous in the sense that they are not determined within the model, they are generally not econometrically exogeneous; that is, they will generally correlate with the market shares s_{jt}. Indeed, we expect a product's market share to increase when the product's unobservable quality increases. This gives rise to the econometric endogeneity problem that I consider in Section 4.

3.3 Quantity-type models

In the model outlined above, consumers care about products' market shares. I call this the share-type or *S-type* model. We may think that consumers could instead care about the overall quantity (i.e. market share times market size) of consumers choosing a particular product. In this case, we replace each s_d^t with $M_t^d s_t$ and s_t^o with $M_t s_t^o$ to obtain the quantity-type or *Q-type* model.

3.4 Multiple equilibria

The existence of equilibrium market shares, i.e. solutions of (3), is ensured by Brouwer's fixed-point theorem when the market share function and f_j function are continuous. There will generally not, however, be a unique vector of equilibrium market shares. This complicates counterfactual analysis because the effect of changing market characteristics depends on which equilibrium is realized under the counterfactual market characteristics.

I aim to avoid this complication by appealing to the fact equilibria in discrete-choice models with network externalities may be, in a sense, locally unique. Suppose we observe market characteristics χ_t and market shares s_t. By the implicit function theorem, there is a unique Δ^J-valued function s^* defined on a neighbourhood X of χ_t such that $s^*(\chi_t) = s_t$ and $\chi \in X$,

$$s^*(\bar{\chi}) - \sigma_t(\delta_t(s^*(\bar{\chi}))|\bar{\chi}) = 0$$

as long as the following matrix is nonsingular:

$$I - D_\delta\sigma_t(\delta_t(s_t)|\chi_t)D_s\delta(s_t). \tag{5}$$

I call the unique s^* function defined in the neighbourhood of a particular χ_t an *equilibrium surface* around (χ_t, s_t). When I say that an equilibrium s_t at market characteristics χ_t is locally unique, I mean that there is a unique equilibrium surface defined around χ_t. In Appendix A, I discuss the condition that (5) must be nonsingular and 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. In general, the effect of changing a market characteristic depends on which equilibrium is selected under the updated market characteristics. When an equilibrium surface exists around an equilibrium, however, we can use the slope of this equilibrium surface with respect to various market characteristics to quantify the effects of marginal counterfactual changes to these market characteristics. This corresponds to assuming that small changes in market characteristics do not lead the equilibrium market shares to jump to another equilibrium surface. Thus, we can coherently define the marginal effect of a market characteristic on an equilibrium outcome as the gradient of the equilibrium surface around a baseline equilibirum of interest with respect this characteristic.

4 Identification

Although discrete-choice models of demand are generally identified with market data alone under certain conditions concerning the availability of suitable IVs (see Berry and Haile 2014), adding network externalities to these models generally introduces a requirement for microdata and additional assumptions. This section characterizes the identification of semi-nonparametric versions of the models in section 3. It begins by discussing the main identification problems and their solutions in the context of a simple parametric model of network externalities before generalizing this discussion to the semi-nonparametric setting.

4.1 Simple model

Consider a market t in which consumers choose whether to use a dating website. Let $y_{it} = \mathbb{1}\{u_{it} \geq 0\}$ denote an indicator for whether consumer i uses the dating website and

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

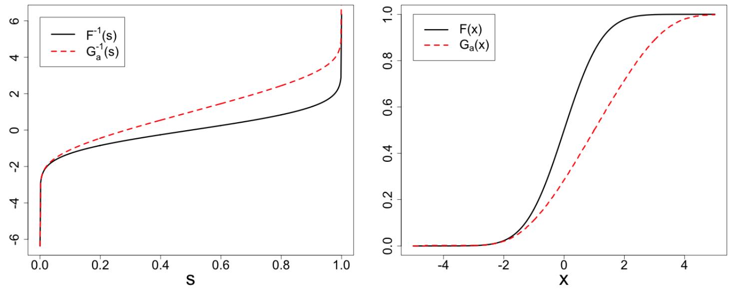
Here, x_t is an exogenous characteristic of the dating website (e.g. its price), s_t is the share of users in market t who use the dating website, and ξ_t is the unobservable quality of the dating website in market t. Additionally, ε_{it} is consumer i's idiosyncratic taste for the dating website; assume that ε_{it} is iid according to the strictly increasing distribution function F. We fix the coefficient of x_t at one as a scale normalization. Integrating over ε_{it} while holding fixed each of s_t, x_t, and ξ_t, we se that the consumer's probability of choosing to use the dating website 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 probability of using the site, we obtain

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

Figure 3: Illustration of G_δ



or, inverting the strictly increasing distribution function F,

$$F^{-1}(s_t) = x_t + \gamma s_t + \xi_t. \tag{6}$$

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 $\tilde{\theta} = (G_\delta, \gamma + \delta)$, where $\delta \geq 0$ and

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

As Figure 3 illustrates, increasing δ increases the variance of G_δ. This dampens the effect of x_t on s_t. This observational equivalence result reflects that we cannot distinguish between the following two explanations for the observed relationship between x_t and s_t:

(i) Consumers dislike spending on dating websites (direct effect of x_t on market shares; this effect is relative large when δ is low)

(ii) Consumers like that lower prices attract more potential matches to dating websites (indirect network externality effect of x_t on market shares; this effect is relatively large when δ is high).

I now make two remarks. First, the source of the identification problem here is that we cannot fix s_t while varying x_t to identify the effect of x_t on consumers' choice probabilities. Second, note that a researcher who ruled out network externalities when $\gamma > 0$ would overstate the effect of x_t on consumers' choice probabilities by attributing the entire positive relationship between x_t and s_t to the direct effect of x_t on choice probabilities.

Microdata including individual consumers' choices and characteristics allows us to identify the binary choice model outlined above. To see why, suppose that the individual characteristic w_{it} that we observe is a component of the idiosyncratic taste term ε_{it}, i.e. $\varepsilon_{it} = w_{it} + \tilde{\varepsilon}_{it}$. In the dating website setting, the consumer characteristic w_{it} could be a measure of (negative) internet speed. Suppose additionally that the remaining unobservable aspect of idiosyncratic tastes $\tilde{\varepsilon}_{it}$ is distributed according to $\tilde{F}_t$ 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 the average taste for the dating website in market t. We now see that, when the t subscript on $\Pr_t$ means "conditional on market t,"

$$\begin{aligned} \Pr(y_{it} = 1 \mid w_{it} = \bar{w}) &= \Pr_t(u_{it} \geq 0 \mid w_{it} = \bar{w}) \\ &= 1 - \Pr_t(\delta_t + \tilde{\varepsilon}_{it} \leq \bar{w}) \end{aligned} \tag{7}$$

The left-hand side is observable and the final expression on the right-hand side is one minus the distribution function of $\delta_t + \varepsilon_{it}$. When w_{it} has a large support, (7) identifies the distribution of $\delta_t + \tilde{\varepsilon}_{it}$. With a location normalization (e.g. $\mathbb{E}_t[\tilde{\varepsilon}_{it}] = 0$), knowledge of this distribution separately identifies δ_t and $\tilde{F}_t$.

The remaining task of identification is the separate identification of the components of average tastes for the website in market t, i.e. δ_t. The primary challenge in completing this task is the fact that ξ_t and s_t are generally correlated since ξ_t partly determines s_t, which must solve the equation (6) in which ξ_t appears. If w_{it}'s distribution F_t^w varies across markets, then functions of this distribution (e.g. its mean) will correlate with s_t. If F_t^w and ξ_t are independent, then such functions of F_t^w will be suitable excluded instruments for s_t. 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 IV argument.

I now provide an example to build intuition for the identification strategy described above. Suppose that, within some market t, consumers with faster internet speeds are more likely to use the dating website. 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. The difference between our prediction of the site's increase in popularity based on within-market variation and the actual increase in popularity therefore identifies the extent of network externalities. To connect this example 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 (7). Increasing average internet speeds in market t in the example is analogous to using cross-market variation in the distribution of w_{it} to identify the effect of market shares on δ_t.

There are several ways that the IV approach described above could fail. First, the approach is threatened if consumers locate based on ξ_t. As an example, consumers interested in using dating websites may both invest in fast internet and locate in regions where dating websites are exogenously better, which would induce a dependency between ξ_t and F_t^w. Second, if the dating website provides especially high service (or high levels of advertising) in regions where it is most appealing because of the region's distribution of individual characteristics w_{it} (e.g. because the average internet speed is faster in those regions), then ξ_t will depend on F_t^w.

Although microdata is necessary to identify a model of the type discussed above in which market shares directly enter consumers' indirect utilities, market data may be sufficient when consumers instead care about the quantity of consumers choosing each of the products on offer. To see why, redefine u_{it} in the above example as

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

where M_t is the population of market t. When there are no network externalities (i.e. $\gamma = 0$), M_t and s_t will be uncorrelated. When there are network externalities, (i.e. $\gamma \neq 0$), M_t and s_t will correlate, which provides identifying power. Indeed, a positive (negative) correlation between population and the market share would suggest $\gamma > 0$ (respectively, $\gamma < 0$).

4.2 Semi-nonparametric model

I now discuss a semi-nonparametric generalization of the identification analysis in Subsection 4.1, relegating some additional results and proofs to Appendix B. Although I maintain the additive separability of both the market-choice unobservables ξ_{jt} and of the idiosyncratic tastes ε_{ijt} throughout this section, I allow consumers' indirect utilities to flexibly depend on market shares and observable product characteristics. I analyze a model without distinct demographic groups in this section for expositional reasons, although I extend the analysis to the case with multiple demographic groups in Appendix B. The particular form of consumers' indirect utilities that I consider is

$$u_{ijt} = \underbrace{g_j(s_t, x_{jt}) + \xi_{jt}}_{=:\delta_{jt}} + \varepsilon_{ijt}. \tag{8}$$

for $j \in \{1, \ldots, J\}$, where the x_{jt} are observable product characteristics. In discrete-choice models of demand, one of the product characteristics x_{jt} is typically price. In what follows, I ignore price endogeneity to focus attention on network externalities. I impose $u_{i0t} = 0$ as a location normalization. The ε_{ijt} deviates capture all individual variation in tastes for j. Note that, in the market-data setting, we do not observe individuals' characteristics so that ε_{ijt} includes taste heterogeneity owing to heterogeneity in consumer characteristics that are observable in the microdata setting. The ε_{ij} are not necessarily independently or identically distributed across goods for a given consumer. From here on, I will typically suppress the market subscript t. I now provide some assumptions.

Assumption ACDIST (Absolutely continuous distribution)**.** The distribution of $\{\varepsilon_{ij}\}_{j\in\mathcal{J}}$ conditional on χ is absolutely continuous with respect to the Lebesgue measure with probability one and admits a density that is positive everywhere on $\mathbb{R}^J$.

Assumption EPS-Market (Restrictions on ε_{ijt} for market-data identification)**.** $\{\varepsilon_{ij}\}_{j=1}^{J}$ is independent of x and ξ.

Assumption ACDIST allows for the inversion of demand in the sense of Berry et al. (2013), as discussed in Appendix B. Assumption EPS-Market allows us to write $\sigma(\delta|\chi)$ as $\sigma(\delta)$. Note that Assumption EPS-Market does not rule out a dependence of ε_{ijt} on all observable product characteristics as occurs in Berry et al. (1995, henceforth BLP)-style random coefficient models; in general, I only require a scalar-valued x_{jt} for my identification results, and we can condition on other product characteristics and suppress them in our notation as in Berry and Haile (2014). In the random coefficients setting, this requires that at least one product characteristic has a fixed coefficient. I implicitly maintain both Assumption ACDIST and EPS-Market throughout this section. Last, I assume all expectations in my analysis exist rather than explicitly state required moment conditions throughout this section.

I first study identification in the market-data setting when the distribution of ε_{ijt} is known. The following results (i) highlight how this questionable assumption can be pivotal for the identification of the model and (ii) suggest how the model may be identified with microdata.

Assumption KD (Known distribution)**.** The random variables $\{\varepsilon_{ij}\}_{j\in\mathcal{J}}$ are distributed according to the distribution function F_ε, which is known to the researcher and satisfies Assumption ACDIST.

Assumption EXCL (Exclusion of choice externalities)**.** For each $j \in \mathcal{J}$, there is an $\ell \in \mathcal{J} \setminus \{j\}$ such that g_j does not depend on s_ℓ; that is, for any $s, \hat{s} \in \Delta^J$ such that $s = (s_1, \ldots, s_\ell, \ldots, s_J)$ and $\hat{s} = (s_1, \ldots, \hat{s}_\ell, \ldots, s_J)$, we have $g_j(s, x_j) = g_j(\hat{s}, x_j)$ for all $x_j \in \operatorname{supp} x_j$.[1]

Note that Assumption KD is invoked, for instance, when the ε_{ij} are assumed to follow a type 1 extreme value distribution. Additionally note that Assumption EXCL cannot hold when $J = 1$, i.e. in the binary choice setting.

Proposition 1. *Suppose that Assumptions KD and EXCL hold. Then, g_j, and ξ_j are identified under completeness for all j.*

See Appendix B.3 for the proof.

Although Proposition 1 concerns identification in the market-data setting, it suggests how the model is identified in the microdata setting. Suppose that the researcher observes individual characteristics $\{w_{ij} : j \in \mathcal{J}\}$ and choices $\{y_{ij} : j \in \mathcal{J}\}$, where

$$y_{ij} = \mathbb{1}\left\{ j = \arg\max_{k \in \mathcal{J}} \ u_{ik} \right\}. \tag{9}$$

Suppose in addition that

$$\varepsilon_{ij} = \lambda w_{ij} + \tilde{\varepsilon}_{ij},$$

where $\{\tilde{\varepsilon}_{ij}\}_{j \in \mathcal{J}}$ follows some distribution $G_{\tilde{\varepsilon}|x}$ conditional on x. I impose $\mathbb{E}[\tilde{\varepsilon}_{ij}|x, s, \xi] = 0$ as a location normalization. When $\lambda = 1$ is our scale normalization, we obtain

$$u_{ij} = w_{ij} + \delta_j + \tilde{\varepsilon}_{ij},$$

where $\delta_j = g_j(s, x_j) + \xi_j$. By results on the identification of discrete-choice models with microdata, δ_j and $G_{\tilde{\varepsilon}|x}$ are identified under a large-support assumption on w_{ij}. I formalize this statement in the subsequent proposition.

Proposition 2. *Suppose* $\operatorname{supp} w_{it}|t = \mathbb{R}^J$, *where conditioning on t means conditioning on x_t, ξ_t, and s_t. Then, $\{\delta_{jt}\}_{j \in \mathcal{J}_t}$ is identified for all t. Furthermore, $G_{\varepsilon|x}$ is identified on the support of x_t.*

See Appendix B.4 for the proof. Given that w_{ij} is observable and independent of $\tilde{\varepsilon}_{ij}$, the identification of the distribution of ε_{ij} immediately follows. With this distribution in hand, we can apply Proposition 1; that is, the model is identified under some mild assumptions as long as $J \geq 2$. In my main discussion of identification with microdata, I describe how to identify the model even in the binary choice setting where $J = 1$.

I now turn to the case in which the econometrician does not know F_ε and instead desires to identify it using market data. The upcoming Proposition 3 tells us that the model is not generally identified in this case; thus, the assumption of a known distribution is pivotal for identification in Proposition 1. Proposition 3's negative result is robust to a special-regressor restriction on g_j that implies a scale normalization and is popular in the literature on the identification of discrete-choice models.

Assumption SREG-Market (Special regressor — market data)**.** For each $j \in \mathcal{J}$, there is a function $h_j : \Delta^J \to \mathbb{R}$ such that

$$g_j(s, x_j) = x_j + h_j(s)$$

[1] I use $\operatorname{supp} z$ to denote the support of a random variable z.

for a scalar-valued product characteristic x_j.

Proposition 3. *Fix some model primitives $\theta \coloneqq (F_\varepsilon,\, g)$, where $g = (g_1, \ldots, g_J)$. For any distribution $\tilde{F}_\varepsilon$ over $\mathbb{R}^J$ satisfying Assumption ACDIST, there are functions $\tilde{g} = (\tilde{g}_1, \ldots, \tilde{g}_J)$ such that $\tilde{\theta} = (\tilde{F}_\varepsilon,\, \tilde{g})$ is observationally equivalent to θ. This remains true under Assumption SREG-Market.*

See Appendix B.5 for a proof.

I now expand my discussion of identification with microdata. Consider the indirect utilities

$$u_{ijt} = g_j\left(w_{ijt},\, s_t,\, x_{jt}\right) + \xi_{jt} + \varepsilon_{ijt}$$

where, for each j, w_{ijt} is an observable, scalar-valued consumer characteristic. Suppose that we also observe each consumer i's choice y_{it} as defined by (9). This model captures heterogeneity in preferences for product characteristics in that consumer characteristics w_{ijt} and product characteristics x_{jt} are permitted to interact in the nonparametric g_j functions and the distribution of ε_{ijt} is permitted to depend on x_{jt}. See Appendix B.1 for a discussion of how the model can be generalized so that indirect utilities even more flexibly depend on additional individual and product characteristics without significantly affecting this section's remaining identification results.

The upcoming Proposition 4 is analogous to Proposition 2 in that it establishes that, under a large support assumption, within-market variation in consumer characteristics identifies products' central utilities δ_{jt} in addition to their distributions of idiosyncratic tastes ε_{ijt}.

Assumption EPS-Micro (Restrictions on ε_{ijt} for microdata identification)**.** The distribution of ε_{ijt} conditional on x_t does not vary across markets and ε_{ijt} is independent of $(\xi_t,\, w_{ijt})$.

Assumption SREG-Micro (Special regressor — microdata)**.** For each $j \in \mathcal{J}_t$, the function g_j takes the form.[2]

$$g_j\left(w_{ijt},\, s_t,\, x_{jt}\right) = w_{ijt} + \tilde{g}_j\left(s_t,\, x_{jt}\right).$$

Assumption LSUPP (Large support)**.** The support of $\{w_{ijt}\}_{j=1}^{J}$ conditional on $\{(x_{jt}, \xi_{jt})\}_{j=1}^{J_t}$ and s_t is $\mathbb{R}^{J_t}$

Proposition 4. *Under Assumptions SREG-Micro and LSUPP, $(F_\varepsilon,\, \delta_t)$ is identified, where*

$$\delta_{jt} = \tilde{g}_j(s_t, x_{jt}) + \xi_{jt}. \tag{10}$$

Proof. The proof is almost identical to that of Proposition 2. □

With δ_{jt} in hand, the identification of $\tilde{g}$ and ξ follow from a straightforward nonparametric IV (NPIV) argument. The upcoming assumptions provide characteristics of instruments z_{jt} that identify the model. After stating Proposition 5, which is the the main microdata identification result of this section, I discuss available instruments that plausibly satisfy these assumptions.

Assumption NPIV-EX (Exclusion restriction for NPIV)**.** The unobservables ξ_{jt} are mean-independent of the instruments z_{jt}, i.e. $\mathbb{E}[\xi_{jt}|z_{jt}] = 0$.

[2]Note that, in Assumption SREG-Micro, the imposition of a coefficient of one on w_{ijt} is a scale normalization.

Assumption NPIV-C (Completeness condition for NPIV)**.** For all functions Δ such that $\mathbb{E}[\Delta(s_t, x_{jt})|z_{jt}] = 0]$ almost surely, $\Delta(s_t, x_{jt}) = 0$ almost surely.

Proposition 5. *Suppose that the conditions of Proposition 4 hold. Additionally suppose that Assumptions NPIV-EX and NPIV-C hold. Then, $\{\tilde{g}_j\}_j$, ξ_t, and F_ε are identified.*

Proof. Each of δ_t and F_ε are identified by Proposition 4. Apply the NPIV argument of Newey and Powell (2003) to (10) with the instruments z_t from Assumptions NPIV-EX and NPIV-C to separately identify $\tilde{g}_j$ and ξ_{jt}. □

My identification argument is nearly identical to that in Berry and Haile (2010) with the exception that I instrument for s_t rather than a product characteristic (e.g. price) correlated with ξ_{jt}. The most promising way to construct the instruments z_{jt} appearing in Proposition 5 is to use cross-market variation in market-specific distributions of consumer characteristics w_{ijt}. To provide some examples, we could set $z_{jt} = \mathbb{E}_t[w_{ijt}]$, the mean value of w_{ijt} in market t, or $z_{jt} = \Pr_t(w_{ijt} + \varepsilon_{ijt} = \max_k w_{ikt} + \varepsilon_{ijt})$, the predicted share of product j in market t when $\delta_{kt} = 0$ for all k. As long as F_t^w is suitably independent of ξ_t, these instruments will satisfy Assumption NPIV-EX. Unless $J_t = 1$, exogenous characteristics of other products are also available as IVs. See Subsection 4.4 for additional discussion of available instruments.

4.3 Quantity-type models

As outlined in Subsection 3.3, network externality models may feature consumers who care about products' market shares or their quantities, i.e. market shares times market sizes. This section considers the identification of a model of the latter type. I begin by considering identification with market data when indirect utilities take the form

$$u_{ijt} = \underbrace{g_j(M_t s_t, x_{jt}) + \xi_{jt}}_{=:\delta_{jt}} + \varepsilon_{ijt},$$

which is identical to (8) with the exception that $M_t s_t$ appears in the place of s_t. Suppose that the distribution of $\varepsilon_{ijt}|t$ is unknown. Although the S-type model is not generally identified in this setting, the Q-type model *is* identified under various sorts of restrictions similar to the "special regressor" assumptions made in the literature on the identification of discrete-choice models. In addition to these functional form restrictions, we require that M_t is exogeneous to ξ_t in a mean-independence sense so that it is a valid IV.

Proposition 6. *Suppose that Assumption ACDIST holds and that*

$$\mathbb{E}[\xi_t|x_t, M_t] = 0. \tag{11}$$

Then, the following list gives identification results under various assumptions on g:

(i) Separability of network externality:

$$g_j(M_t s_t, x_{jt}) = \alpha M_t s_{jt} + \tilde{g}_{jt}(x_{jt}),$$

for all $j \in \mathcal{J}_t$ and for $\alpha \in \{-1, 1\}$, where α is known by the researcher. In this case, g, ξ, and F are identified.

(ii) *Special regressor assumption:*

$$g_j(M_t s_t, x_{jt}) = x_{jt} + \tilde{g}_j(M_t s_t)$$

for all $j \in \mathcal{J}_t$. In this case, ξ_t and $\kappa_j(s_t, M_t) := \sigma_j^{-1}(s_t) - \tilde{g}_j(M_t s_t)$ are identified. Under either of the following assumptions, $\tilde{g}$ and σ are separately identified:

(a) *Own shares: $\tilde{g}_j(s)$ depends only on s_j for all $j \in \mathcal{J}_t$.*

(b) *Exchangeability and exclusion: Let $\rho^j(s)$ be some vector constructed by permuting the components of s while fixing the position of the jth component so that $s_j = \rho_j^j(s)$. We assume that $\sigma_j^{-1}(s) = \sigma_j^{-1}(\rho^j(s))$. Additionally, assume that for each $j \in \mathcal{J}_t$, there is an $\ell \in \mathcal{J}_t$ such that $\tilde{g}_j$ does not depend on its ℓth argument.*

Proof. See Appendix B.6. □

The analysis of the identification of Q-type models with microdata follows almost exactly as my analysis of the S-type model with microdata in Subsection 4.2. The primary differences are that (i) s_t is replaced my $M_t s_t$ in the consumer's indirect utilities and (ii) M_t becomes available as an IV when we assume that it satisfies the mean-independence exclusion restriction.

4.4 Discussion of identification results

I now discuss my identification results and relate them to the literature on discrete-choice models with network externalities. As mentioned in Subsection 1.1, Bayer and Timmins (2007) consider the estimation of an S-type network externality model with microdata using an IV approach. They require IVs because the market shares and the demand unobservables appearing in consumers' utilities are generally correlated with each other. Bayer and Timmins (2007) and Timmins and Murdock (2007) address this problem by using transformations of other products' observable characteristics as excluded instruments; henceforth, I refer to such instruments as *BLP instruments* since they are used as instruments for prices in BLP and much of the pursuant literature. This approach is suitable for the Timmins and Murdock's (2007) empirical application, in which there is a single market and therefore no possibility of using instruments that exploit cross-market variation. My identification analysis formally establishes the validity of Bayer and Timmins's and Timmins and Murdock's approaches under the exogeneity and completeness conditions on the BLP instruments. Additionally, my analysis shows how Timmins and Murdock (2007) could obtain additional instruments if they used data from multiple markets in their analysis.

My identification analysis also relates to the IVs used by Guiteras et al. (2019, hereafter GLM). GLM model demand for latrines in Bangladeshi villages using a discrete-choice model with network externalities and estimate this model with data from a randomized control trial. Their experiment provided latrine subsidies to a subset of households in their sample with randomization in the provision of subsidies at household, neighbourhood, and village levels (i.e. the share of households eligible for subsidies varied across neighbourhoods within a village and across villages). They propose a discrete-choice demand model with network externalities and use the share of households eligible for subsidies in a neighbourhood as an instrument for the share of households buying latrines in that neighbourhood. The instruments used by GLM are examples of the instruments based on market-specific distributions of consumer characteristics that I consider in my identification analysis. My paper expands upon GLM's identification insight by showing that demographic characteristics not directly related to a characteristic of the analyzed product can aid in identifying network externalities. I also note that, in a setting with more than two choices, GLM-type instruments can be combined with BLP instruments in identifying network externalities.

Both the instruments used by GLM and my own instruments are analogous to the price instruments used by Waldfogel, which are transformations of market-specific distributions of consumer characteristics. These instruments are relevant in the price-endogeneity setting because firms set market-specific prices based on local demand, which in turn depends on the characteristics of consumers in that market. They are relevant in the network externality setting because heterogeneity in market-level distributions of consumer characteristics induce variation in products' popularity across markets. Following Berry and Haile (2016), who formally analyze these instruments in the price-endogeneity setting, I call instruments computed from market-level distributions of consumer characteristics *Waldfogel instruments*. As noted in Subsection 4.1, there are several ways that the Waldfogel instruments could fail to satisfy the exclusion restriction in the dating website application. These include (i) the possibility that firms provide higher quality ξ_{jt} in markets with favourable demographic profiles and (ii) consumers choose their geographical markets based on the local popularity of their preferred products. To illustrate this second point, consider a setting with just one "inside" product that is favoured by younger consumers. Suppose that young consumers, motivated by their desire to consume the product, move to markets where the product's quality ξ_t (unobserved to the econometrician) is high. This behaviour induces a correlation between markets' ξ_t values and the shares of their populations that are young. When age is the consumer characteristic w_{it}, this correlation invalidates the Waldfogel instruments.

The relevance of the Waldfogel instruments becomes weaker as we specify our demographic groups more narrowly. This is because the Waldfogel instruments rely on variation in consumer characteristics within subpopulations of markets belonging to particular demographic groups to shift products' market shares within these subpopulations. Defining demographic groups more narrowly means restricting the variation in consumer characteristics within these subpopulations, i.e. reducing the variation of the Waldfogel instruments. To fix ideas, consider the extreme case in which there are two demographic groups, call them young and old, each of which has a distinct mean taste for a particular product. The proportion of a market that is young shifts the product's market share across markets. But, when the only market shares that the consumer cares about are those within her own age group, the proportion of a market that is young does not shift the market share of the product among consumers that are young (or among consumers that are old). Therefore, the Waldfogel instrument computed as the proportion of young consumers in the market becomes irrelevant when we move from a model with only one demographic group to one in which the demographic groups are age groups.

5 Estimation

Estimation of the model proceeds in two steps. In the first step, which I call the *microstep*, I estimate mean tastes for each site in each market among each demographic group in addition to the contribution of consumer characteristics to tastes for particular sites. In the second step, which I call the *market step*, I estimate the contribution of site characteristics and market shares to mean tastes for sites.

5.1 Microstep

The estimating equation of the microstep is

$$u_{ijt} = \delta_{jt}^{d(i)} + w'_{ijt}\lambda + \varepsilon_{ijt}, \tag{12}$$

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 (12) via maximum likelihood. Note that the microstep is the same for each of the S-type and Q-type models.

5.2 Market step

The estimating equation of the market step of the S-type model is

$$\delta_{jt}^d = x'_{jt}\beta + f_j^d\left(s_t^o, s_t^1, \ldots, s_t^D; \gamma\right) + \xi_{jt}^d. \tag{13}$$

See Subsection 3.2 for a description of the terms appearing in this equation. The f_j^d function, which I call the network externality function, is known up to the finite-dimensional parameter vector γ. Given that δ_{jt}^d is not directly observed, I substitute estimates of this quantity from the microstep for the true quantity in the actual estimation routine. I similarly substitute estimates of the market shares computed using the Comscore data in for the true market shares. I choose the x_{jt} to be firm-time period indicators and denote their coefficients by $\psi_{j\tau}$ so that $x'_{jt}\beta = \psi_{j\tau}$, where τ is the time period to which market t belongs (recall that markets are defined by both geography and time period). Last, I estimate (13) using several different parmetric forms of the network externality function $f_j^d(\cdot;\gamma)$, as I state explicitly in Section 6

I use excluded instruments to consistently estimate f_j^d given that market shares generally correlate with ξ_{jt}^d. For each market share appearing in (13), I included as an instrument the predicted value of the market share based on the microstep estimates when $\delta_{jt}^d = 0$ for all j, t, d. 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 (13). To illustrate, suppose $D = 1$ and the specific estimating equation is

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

Then, the excluded instrument I use is

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

where M_t is the sample size of consumers in market t. I similarly construct instruments for other specifications.

The market step of estimating a Q-type model proceeds similarly. For the Q-type model, quantities replace market shares in (13) and predicted quantities are used in constructing instruments instead of predicted market shares.

5.3 Price sensitivity

To this point, I have ignored price competition between firms. Estimating consumers' 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 my sample charge uniform prices across geography means that I observe minimal price variation, which prevents me from estimating price sensitivity alongside other parameters of the consumer choice model. 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_{jt} in (13) and let $\psi_{j\tau}$ denote the fixed effect for site j in time period τ. I then make the decomposition $\psi_{j\tau} = \bar{\psi} - \alpha p_{j\tau}$ and assume that the observed prices $\{p_j^*\}$ constitute a Bertrand-Nash pricing equilibrium with marginal costs of zero in that

$$p_j^* = \arg\max_{p_j} \sum_t M_t \sigma_{jt}(p_j, p_{-j}^*; \alpha) p_j \qquad \forall j \text{ s.t. } p_j^* > 0. \tag{14}$$

The profit maximization problem in (14) gives rise to the first-order conditions (FOCs)

$$\sum_t \left[M_t \frac{\partial \sigma_{jt}}{\partial p_j}(p_j^*, p_{-j}^*; \alpha) p_j^* + \sigma_{jt}(p_j^*, p_{-j}^*; \alpha) \right] = 0 \tag{15}$$

which provided the basis of my estimation of α.

To compute my estimator $\hat{\alpha}$ of α, I substitute empirical analogues/estimates for population objects/parameters in each paid site's FOC (15) and then solve for α (there turns out to be an explicit closed-form when we assume ε_{ijt} is iid type 1 extreme value.) These FOCs include price derivatives of firms' market shares, which are not well defined without an assumption on how prices affect equilibrium selection in the presence of multiple equilbria. 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 Subsection 3.4, 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 used in my estimation of α. The implicit function theorem provides an explicit form for these derivatives. Each derivative $\partial\sigma_{jt}/\partial p_j$ reflects two effects of price on market shares: a *direct* effect of price on consumers' likelihoods of purchasing product j and an *indirect* effect owing to the fact that the direct effect changes good j's market share which in turn changes the network externality term in good j's indirect utility. Some of the sites in my sample are free to use (i.e. their prices are zero). 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.

One way to check whether $\hat{\alpha}$ is a reasonable estimator is to compute the profit-maximizing price for a monopolist and confirm that this price is sensible. Subsection 7.4 considers a counterfactual in which `match.com` becomes a monopolist. In this counterfactual, `match.com` raises its price by 12.9% (see Table 18); this magnitude seems sensible.

6 Parameter estimates

This section reports and discusses my parameter estimates. I estimate the model using several different specifications of demographic groups. In particular, the specifications that I consider are

(**Overall**) All consumers belong to the same demographic group

(**Race**) Non-black consumers belong to the first demographic group and black consumers belong to the second demographic group.

(**Age**) 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 Subsection 5.2. The sites that consumers choose between are `eharmony.com`, `match.com`, `okcupid.com` and `pof.com` in the 2007–2008 time period and each of these sites in addition to `zoosk.com` in the 2017–2018 time period; choices to use other sites and a failure to use any site are grouped together in the outside option.

6.1 Microstep

Table 6 displays estimates from the microstep for the "Overall" specification of demographic groups. In particular, the table reports estimates of the λ parameters governing the effects of demographic characteristics on consumers' tastes for particular sites. Although many of the parameter estimates are not statistically significant at the 95% level, several estimates indicate significant taste differences across demographic groups. Black consumers, for instance, dislike `match.com`, `okcupid.com` and `pof.com` relative to white consumers. Consumers with broadband internet generally like dating websites more than consumers without broadband. Consumers with higher incomes have lower tastes for dating websites, especially for `eharmony.com` and `pof.com`.

The microstep estimates for the "Race" and "Age" demographic group specification are included in Appendix D. The results for these alternative specifications are qualitatively similar to those reported in Table 6.

6.2 Market step

Panel A of Table 7 displays the parameter estimates of an S-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 fact that the unobservables ξ_{jt} and market shares s_{jt} are positively correlated. Panel A also reports my estimate of α for this specification. Note that the rows with names of dating websites (e.g. "eharmony: 1") provide the estimated site-time period intercepts for the first (2007–2008) time period; the estimated intercepts for the second time period are omitted.

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 $\tilde{\zeta}_{jt}$

Table 6: First-stage parameter estimates – “Overall” demographic group specification

	`eharmony.com`	`match.com`	`okcupid.com`	`pof.com`	`zoosk.com`
Education: High school or less (Omit.)	0.000	0.000	0.000	0.000	0.000
	-	-	-	-	-
Education: Some college	-0.119	0.009	0.099	0.110	-0.231
	(0.119)	(0.072)	(0.281)	(0.185)	(0.514)
Education: College degree	-0.104	-0.003	0.354	0.075	-0.216
	(0.123)	(0.074)	(0.266)	(0.185)	(0.492)
Education: Advanced degree	0.051	0.071	-0.337	-0.267	-0.570
	(0.140)	(0.086)	(0.415)	(0.266)	(1.102)
Education: Unknown	0.046	-0.183	-0.317	0.194	-1.443
	(0.086)	(0.055)	(0.224)	(0.146)	(0.551)
Age: Under 25yo (Omit.)	0.000	0.000	0.000	0.000	0.000
	-	-	-	-	-
Age: 25-29yo	0.010	0.066	-0.017	0.193	2.510
	(0.141)	(0.096)	(0.361)	(0.216)	(3.705)
Age: 30-34yo	-0.094	0.064	-0.339	0.346	3.772
	(0.133)	(0.089)	(0.348)	(0.199)	(3.597)
Age: 35-39yo	0.018	0.033	-0.257	0.187	3.400
	(0.129)	(0.088)	(0.334)	(0.199)	(3.604)
Age: 40-49yo	0.004	-0.003	-0.136	-0.035	2.987
	(0.122)	(0.083)	(0.307)	(0.189)	(3.587)
Age: 50+yo	-0.045	-0.003	-0.282	0.060	3.542
	(0.121)	(0.082)	(0.305)	(0.186)	(3.576)
Children in HH: No (Omit.)	0.000	0.000	0.000	0.000	0.000
	-	-	-	-	-
Children in HH: Yes	0.022	0.000	0.119	-0.049	-0.084
	(0.052)	(0.035)	(0.149)	(0.082)	(0.364)
Race: White (Omit.)	0.000	0.000	0.000	0.000	0.000
	-	-	-	-	-
Race: Black	0.005	-0.493	-0.622	-0.424	-0.189
	(0.066)	(0.053)	(0.234)	(0.119)	(0.426)
Race: Asian	-0.136	-0.335	-0.173	-0.791	0.298
	(0.149)	(0.096)	(0.297)	(0.260)	(0.397)
Race: Other	-0.232	-0.322	0.021	-0.262	0.235
	(0.142)	(0.088)	(0.258)	(0.168)	(0.388)
Broadband: No (Omit.)	0.000	0.000	0.000	0.000	0.000
	-	-	-	-	-
Broadband: Yes	0.534	1.098	1.218	0.250	2.140
	(0.081)	(0.068)	(0.338)	(0.118)	(3.414)
Hispanic: No (Omit.)	0.000	0.000	0.000	0.000	0.000
	-	-	-	-	-
Hispanic: Yes	-0.368	-0.277	-0.564	-0.456	-0.534
	(0.045)	(0.029)	(0.136)	(0.075)	(0.356)
Income: Under 25k (Omit.)	0.000	0.000	0.000	0.000	0.000
	-	-	-	-	-
Income: 25-75k	-0.093	-0.118	-0.107	-0.013	-0.015
	(0.050)	(0.034)	(0.146)	(0.079)	(0.355)
Income: 75-100k	-0.115	-0.161	-0.126	-0.145	-0.457
	(0.063)	(0.042)	(0.181)	(0.102)	(0.540)
Income: Over 100k	-0.242	-0.175	-0.097	-0.412	-0.440
	(0.061)	(0.040)	(0.168)	(0.102)	(0.440)
HH size: 1 (Omit.)	0.000	0.000	0.000	0.000	0.000
	-	-	-	-	-
HH size: 2	-0.022	-0.006	-0.162	-0.008	-0.245
	(0.076)	(0.051)	(0.200)	(0.114)	(0.400)
HH size: 3	-0.155	-0.079	-0.294	0.007	-0.294
	(0.091)	(0.061)	(0.243)	(0.137)	(0.500)
HH size: Over 3	-0.052	-0.096	-0.282	-0.052	-0.573
	(0.089)	(0.060)	(0.236)	(0.134)	(0.489)

Table 7: Market step parameter estimates – "Overall" demographic group specification

Panel A: Parameter estimates

	OLS	IV
$\log(s_{jt})$	0.99	0.77
	(0.01)	(0.07)
eharmony: 1	-0.12	-0.84
	(0.03)	(0.21)
match: 1	-0.48	-1.01
	(0.03)	(0.15)
okcupid: 1	-0.21	-1.43
	(0.06)	(0.36)
pof: 1	-0.06	-1.02
	(0.05)	(0.28)
p_j ($\hat{\alpha}$)		0.0065
		(0.0015)

Panel B: First stage of IV regression

	$\widetilde{\log(s_{jt})}$
$\tilde{z}_{jt}$	4.24
	(0.98)
F	18.74

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, indicating that my instruments are highly relevant.

Tables 8 and 9 report estimates from the "Race" and "Age" demographic group specifications under various specifications of the network externality function. In particular, column (1) of each table reports estimates from 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 consumers care about 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 consumers care about the market share of a site within their own demographic group only, but the demographic groups have different preferences for their own-group market shares; and column (4) reports estimates from a specification in which consumers care about 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 both age and race, as consumer care about market shares within their own demographic group much more than shares within the other group. In each table, 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.

Table 8: Market step parameter estimates – "Race" demographic group specification

	(1)	(2)	(3)	(4)
Own-group $\log\left(s_{jt}^{d}\right)$	0.769	0.777	-	0.728
	(0.128)	(0.166)	-	(0.236)
Own-group $\log\left(s_{jt}^{d'}\right)$	-	0.020	-	-
	-	(0.166)	-	-
Own-group $\log\left(s_{jt}^{\text{nonblack}}\right)$	-	-	0.774	-
	-	-	(0.152)	-
Own-group $\log\left(s_{jt}^{\text{black}}\right)$	-	-	0.764	-
	-	-	(0.203)	-
$\log\left(s_{jt}^{\text{overall}}\right)$	-	-	-	0.092
	-	-	-	(0.293)
p_{jt} $(\hat{\alpha})$	0.0064	0.0041	0.0063	0.0052
	(0.0036)	(0.0048)	(0.0084)	(0.0094)

Table 9: Market step parameter estimates – "Age" demographic group specification

	(1)	(2)	(3)	(4)
Own-group $\log\left(s_{jt}^{d}\right)$	0.781	1.215	-	-2.783
	(0.055)	(0.511)	-	(55.125)
Own-group $\log\left(s_{jt}^{d'}\right)$	-	-0.398	-	-
	-	(0.511)	-	-
Own-group $\log\left(s_{jt}^{\text{younger}}\right)$	-	-	0.796	-
	-	-	(0.075)	-
Own-group $\log\left(s_{jt}^{\text{older}}\right)$	-	-	0.770	-
	-	-	(0.075)	-
$\log\left(s_{jt}^{\text{overall}}\right)$	-	-	-	4.163
	-	-	-	(64.453)
p_{jt} $(\hat{\alpha})$	0.0009	0.0008	0.0010	-0.0008
	(0.0002)	(0.0004)	(0.0003)	(0.0293)

Table 10: Market step parameter estimates – "Overall" demographic group specification, Q-type model

Panel A: Parameter estimates

	OLS	IV
$\log(M_t s_{jt})$	1.00	0.78
	(0.01)	(0.06)
$\log(M_t)$	-0.97	-0.76
	(0.01)	(0.06)
eharmony: 1	-0.52	-1.04
	(0.07)	(0.21)
match: 1	-0.88	-1.22
	(0.07)	(0.17)
okcupid: 1	-0.59	-1.60
	(0.08)	(0.33)
pof: 1	-0.45	-1.20
	(0.07)	(0.26)
p_j $(\hat{\alpha})$		0.0062
		(0.0016)

Panel B: First stage of IV regression

	$\widetilde{\log(s_{jt})}$
$\tilde{z}_{jt}$	4.12
	(0.99)
F	17.43

6.3 Value of network externalities

We can use the results presented in the previous section to compute the value of an increase in a site's usership to an inframarginal of the site in dollar terms. This is a more readily interpretable measure of the magnitude of network externalities than the estimated parameters of the network externality function. The estimates for the S-type model presented in Table 7 imply that a 10% increase in the usership of a site is worth \$11.70 a month to an inframarginal user of that site whereas the estimates for the Q-type model presented in Table 10 imply that a 10% increase in the usership of a site is worth \$12.60 a month to an inframarginal user of that site. Given that `match.com`'s monthly subscription price of in 2007 was \$34.99, a 10% increasing in `match.com`'s usership in a particular market would be worth 33% of the site's price to an inframarginal `match.com` user under the S-type model estimates and 36% under the Q-type model estimates.

7 Counterfactual analysis

I now use my model estimates to assess the effects of changes in the market environment. The model estimates that I use in this counterfactual analysis are those for the S-type model without distinct demographic groups as presented in Tables 6 (microstep estimates) and 7 (market step estimates). Additionally, I conduct all of my counterfactual analysis for the 2007–2008 time period. By the "market environment," I mean the characteristics of sites competing in the dating website industry, the distribution of consumer characteristics in the analyzed markets, and the parameters governing consumer tastes. I compare outcomes in equilibria under counterfactual market environments with outcomes in equilibria under the baseline market environment, i.e. the observed characteristics of markets/products and the estimated taste parameters. The baseline equilibria to which I compare counterfactual equilibria feature prices and market shares different than those that I observe because neither of the paid sites maximizes its profits under the observed prices.[3] I compute the equilibria described in this section using a *nested best response algorithm.* This algorithm is nested in the sense that, first, it iterates on market shares to find a fixed-point of the market share function for a particular vector of prices. It then iterates on prices to find a price vector such that each firm's price maximizes its own profits given the other firm's price.

This section conducts several welfare comparisons between counterfactual and baseline equilibria. To compare welfare between equilbria, I compute the average expected utility of a consumer in a market under each of the equilbria and compute the difference between the welfare in the counterfactual equilibrium and the baseline equilibrium. The price-sensitivity estimate $\hat{\alpha}$ enables a conversion of this difference in utils terms to one in dollars.

The first counterfactual analysis that I consider facilitates a decomposition of cross-market variation in sites' market shares into parts depending on network externalities and on taste heterogeneity. Next, I marginally adjust the characteristics of some dating websites to assess the effects of the emergence of a niche site and of increased concentration in the market for dating websites. As noted by Subsection 3.4, it is generally difficult to compute the effects of counterfactual changes in market characteristics on outcomes like prices or welfare when there

[3]Recall that my price sensitivity parameter estimate $\hat{\alpha}$ is the average of the price sensitivity parameter value that justifies `eharmony.com`'s observed price as profit-maximizing and the value that justifies `match.com`'s observed price as profit-maximizing. Thus, the value $\hat{\alpha}$ justifies neither firm's observed price.

is a multiplicity of equilibria. This is because these effects will depend on which equilibrium is selected under the counterfactual market characteristics. I argued in Subsection 3.4 that the concept of local uniqueness of equilibria allows us to define a coherent concept of marginal effects of a market characteristic on outcomes. These marginal effects are the gradients of equilibrium surfaces (defined in Subsection 3.4) with respect to market characteristics. In practice, I assess the effects of counterfactual changes by, first, computing equilibria after adjusting market characteristics. I check that the counterfactual equilibrium market shares I obtain are close to the baseline market shares and conclude that they fall on the same equilibrium surface as the baseline market shares if this is indeed the case. Last, I interpret differences between outcomes in counterfactual equilibria and baseline equilibria as the marginal effects of market chararacteristics on these outcomes.

I also study the effects of increasing concentration in the dating website market by counterfactually establishing a monopoly in the market with `match.com` as the only site. The problem of multiplicity of equilibria under network externalities arises because consumers who enjoy sharing a platform with other consumers may cluster on one dominant site whose identity is indeterminate (i.e. which site is most popular may differ across equilibria). This problem becomes less serious when there is just one site. In fact, I have not found multiple equilibria in a counterfactual market environment in which `match.com` is a monopolist.

7.1 Decomposition of variance in market shares

In this section, I decompose the variation in websites' market shares across geographical markets into three sources: network externalities; unobserved differences in taste across markets; and observed differences in tastes across markets reflecting demographic differences. Table 11 displays the results of this variation decomposition. For each regime, I compute the standard deviation of firms' market shares across geographies for the time period 2007–2008. I subtract sites' average shares across markets from their market shares to ensure that the reported standard deviations capture variation across markets and not variation across sites. For the "No agglom." column, I remove the network externality function from consumers' indirect utilities and re-compute market shares in each market.[4] For the "No ξ_{jt}" column, I set all of the market-level mean unobserved taste unobservables ξ_{jt} to zero. For the "No demo." column, I compute each market's market shares using the distribution of demographic characteristics across all markets in my sample instead of the market's own distribution of demographic characteristics.

Table 11 tells us that network externalities explain most of the variation in market shares across geographical markets; removing market shares from consumers' indirect utilities reduces the cross-geography market share standard deviation by 58%. The table also tells us 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; their contribution to variation in market shares across geographies is crucial for my instrumental variables approach to the identification of network externalities.

[4] Note that eliminating the network externality function from consumers' indirect utilities changes the mean quality of each site and hence changes the market share of the outside good. Thus, I add a constant $f^{\dagger}$ to each site's mean utility δ_{jt}, where $f^{\dagger}$ is chosen so that the market share of the outside good under the model without network externalities is equal to the observed market share of the outside good.

Table 11: Decomposition of geographic variation in market shares

	No agglom.	No ξ_{jt}	No demo.
Share of baseline SD	0.42	0.14	0
Δ	0.58	0.28	0.14

Table 12: Market shares in top markets under the baseline equilibrium

Market	`match.com`	`eharmony.com`
Los Angeles (CA)	0.070	0.050
New York (NY)	0.084	0.066
Chicago (IL)	0.059	0.069
New York (NJ)	0.080	0.062
San Jose (CA)	0.071	0.061
Miami (FL)	0.093	0.065
Dallas (TX)	0.069	0.056
Atlanta (GA)	0.067	0.077

7.2 Increase in market concentration counterfactual

I now assess the effects of increasing market concentration on prices and welfare. In particular, I increase the value of `match.com`, the most popular site in my data, by \$2.00 for all consumers while decreasing the value of `eharmony.com`, the second most popular site, by the same amount for all consumers. This amounts to increasing all of the δ_{jt} parameters for `match.com` by $2 \times \hat{\alpha}$ and decreasing all of the δ_{jt} parameters for `eharmony.com` by the same amount. I focus on welfare in the eight most populous markets in my data; Table 12 gives the market shares of the two most popular dating websites, i.e. `eharmony.com` and `match.com`, under the baseline pricing equilibrium. Note that "New York (NY)" and New York (NJ)" appear as distinct markets because I intersect CSA boundaries with state boundaries when defining my markets; see Section 2 for details.

Table 13 reports the distribution of counterfactual market changes. In particular, it displays various quantiles of the distribution of sites' market-share changes (in percentage points) taken across markets. This table shows that sites' market shares change only marginally in response to the small counterfactual changes in sites' values to consumers, which suggests the counterfactual equilibria belong to the same equilibria surfaces as the baseline equilibria. Table 14 displays the changes in sites' prices in the counterfactual whereas Figure 4 shows the welfare changes (in dollars) for several of the largest markets in the analysis. I decompose the overall welfare change into several sources. Each "Price response" bar gives the average difference between a counterfactual equilibrium in which sites are allowed to adjust their prices from the baseline equilibrium and one in which their prices are held fixed at those from the baseline equilibrium. Each "Network externalities" bar provides the average change in the network externality term in consumers' utilities. Each "Parameter changes" bar provides the average change in consumers' utilities coming directly from the counterfactual change in the δ_{jt} parameters. Last, the "Net change" bar gives the overall average welfare change from the baseline equilibrium to the counterfactual.

Figure 4 and Table 12 show that, among the most populous markets, the net welfare change

Table 13: Market share changes in "increase concentration" counterfactual

	Quantiles of market share changes (%)				
Site	0.01	0.25	0.50	0.75	0.99
eharmony.com	-0.3741	-0.2828	-0.2828	-0.2399	-0.1948
match.com	0.2078	0.2540	0.2540	0.3070	0.3757
okcupid.com	-0.0028	-0.0014	-0.0014	-0.0000	0.0007
pof.com	-0.0075	-0.0033	-0.0033	0.0005	0.0034

Table 14: Price changes in "increase concentration" counterfactual

Site	Baseline	Counterfactual	Difference
eharmony.com	44.36	43.93	-0.43
match.com	46.00	46.50	0.50

from increasing market concentration is positive in the markets where `match.com` commands a larger market share than `eharmony.com` and negative elsewhere. This is in large part because strengthening the national leader `match.com` increases utility from network externalities in markets where this site is also a local leader and decreases utility from network externalities elsewhere. In Subsection 7.1, I found that much of the variation in sites' market shares across markets owes to network externalities. Thus, although network externalities imply greater market concentration for a leading firm can be welfare-increasing, they also induce variation in the identity of the "leading" across geography, which in turn implies heterogeneity across geography in the effects of boosting the national market leader. Note also that the fact that many markets gain from increased concentration reflects that I increase the utilities of inframarginal consumers of `match.com`, the most popular dating website, in my counterfactual exercise. The effect of increase is reflected by the "Parameter change" bars in Figure 4. This part of the welfare decomposition accounts for the lion's share of the "Net change" bar when the net change is positive.

7.3 Emergence of a niche site counterfactual

Network externalities imply that the emergence of a site appealing to a subpopulation of consumers can harm inframarginal users of a popular website with broad appeal by drawing away members of this subpopulation away from the popular site. I now consider a counterfactual that assesses whether this harm is enough to make the emergence of a niche site undesirable. This counterfactual involves marginally increasing (by \$10) the appeal of `pof.com` consumers with college degrees and advanced degrees.

Table 15, Table 16, and Figure 5 are the analogues of Table 13, Table 14, and Figure 4 for what I call the "niche site" counterfactual. Increasing the appeal of `pof.com` increases each component of average welfare. Note that the utility from network externalities increase because the inframarginal users of `pof.com` benefit more from the new users who join the site (many of whom did not use any dating website before) than the inframarginal users of the other sites hurt from these sites' decreases in usership upon the increase in `pof.com`'s appeal. This increase in welfare from network externalities is compounded by an increase in utility from the leading sites' price responses; `pof.com`'s increase in appeal to highly

Figure 4: Welfare changes in "increase concentration" counterfactual

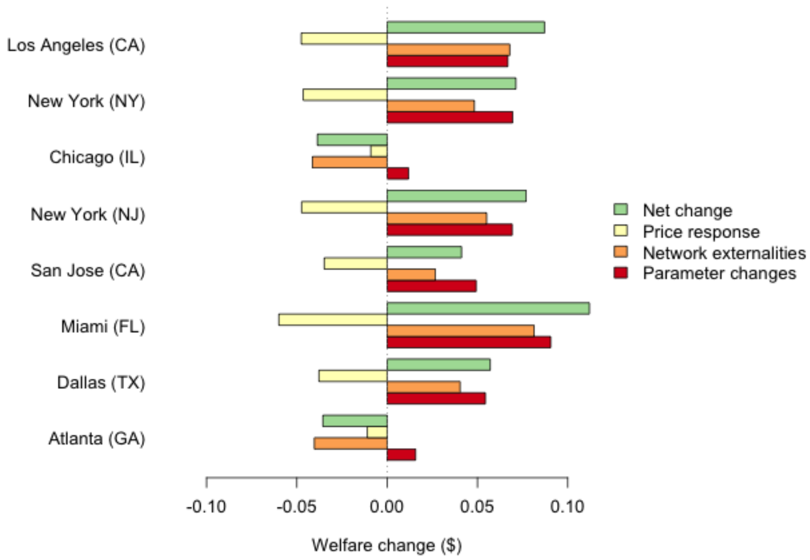


Table 15: Market share changes in "niche site" counterfactual

	Quantiles of market share changes				
Site	0.01	0.25	0.50	0.75	0.99
eharmony.com	-0.0069	-0.0041	-0.0041	-0.0024	-0.0010
match.com	-0.0094	-0.0048	-0.0048	-0.0027	-0.0009
okcupid.com	-0.0009	-0.0005	-0.0005	-0.0003	-0.0001
pof.com	0.0096	0.0181	0.0181	0.0276	0.0422

educated consumers makes the market for dating websites more competitive and induces `eharmony.com` and `match.com` to lower their prices.

7.4 Monopoly counterfactual

I now consider a counterfactual in which `match.com` becomes a monopolist upon the nation-wide exit of all other dating websites. Table 17 provides the changes' in sites' market shares in this counterfactual whereas Table 18 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. Last, Figure 6 provides welfare changes associated with the counterfactual monopolization in the most populous markets. Each "Price response" bar provides the difference in average welfare

Table 16: Price changes in "niche site" counterfactual (%)

	Baseline	Counterfactual	Difference
eharmony.com	44.36	44.35	-0.01
match.com	46.00	45.99	-0.01

Figure 5: Welfare changes in “niche site” counterfactual

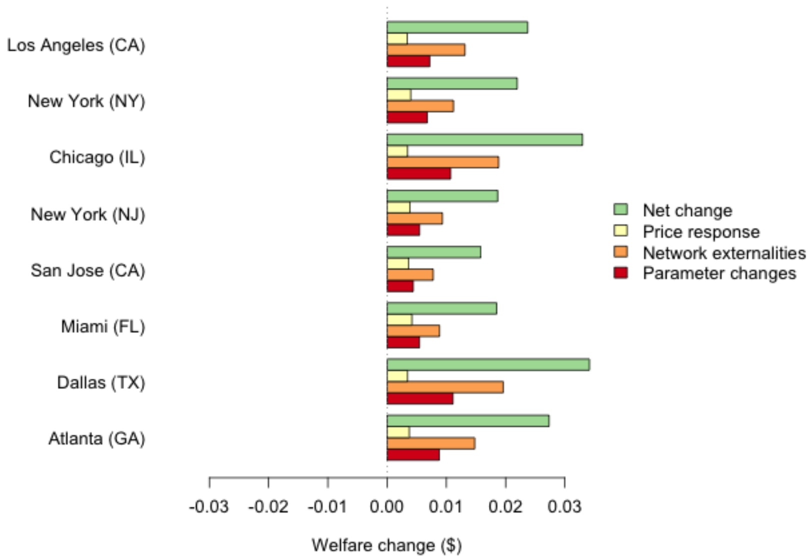


Table 17: Market share changes in the monopoly counterfactual

	eharmony.com	eharmony.com	okcupid.com	pof.com	Outside option
Baseline	0.065	0.073	0.004	0.014	0.844
Counterfactual	0.000	0.087	0.000	0.000	0.913

in a market between a counterfactual equilibrium in which match.com can adjust its price and one in which its price is fixed at its pre-monopoly level. Each “Network externality” bar provides a market’s change in average welfare owing to changes in the utility from network externalities enjoyed by inframarginal consumers of match.com. Last, each “Residual” bar provides the part of the change in average welfare not accounted for by the price response or by the network-externality-related change in welfare experienced by match.com’s inframarginal users. This bar accounts for losses in product variety associated with the monopolization of the market.

Figure 6 tells us that the monopolization of the dating website industry would decrease average welfare by just over \$10/month across the most populous markets. This loss mostly occurs due to a loss in the variety of dating websites available to users. Consumers have heterogeneous tastes for the available sites owing to heterogeneity in these consumers’ observable characteristics (i.e. their w_{ijt} characteristics) and their unobservable idiosyncratic tastes (i.e. their ε_{ijt} shocks); thus, removing sites destroys welfare enhancing matches between consumers and sites. Although inframarginal users of match.com benefit from the site’s larger user base post-monopolization, this benefit is mostly counteracted by match.com’s price increase upon becoming a monopolist.

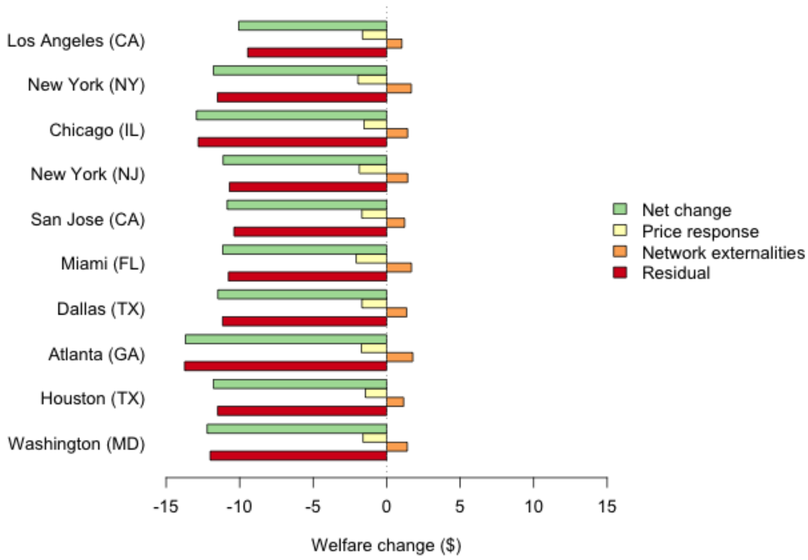
8 Conclusion

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

Table 18: Price change for `match.com` in the monopoly counterfactual

Quantity	Value
Baseline	\$46.00
Counterfactual	\$51.93
Change	12.9%

Figure 6: Welfare changes in the monopoly counterfactual



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 United States. But network externalities do not reverse the general principle that competition benefits consumers. Although heightened concentration increases the benefit of network externalities to inframarginal users of dominant sites, it effects increases in prices and decreases in product variety that significantly reduce consumer welfare.

As noted in the introduction, network externalities arise in many choice settings studied by economists. The model developed in this paper can be straightforwardly adapted to many of these settings, e.g. to demand settings in which fashionability or social pressure is relevant, to settings with congestion, and to residential/educational choice settings where homophily based on social class or race is relevant. The model can also be applied to settings in which consumers prefer to choose platforms with more users because larger user bases lead to more software development on these platforms or to more publicly available expertise about these platforms.

Works cited

Allende, C. 2019. Competition under social interactions and the design of education policies. Working paper.

Bayer, P., R. McMillan, K. Rueben. 2004. An equilibrium model of sorting in an urban housing market. NBER working paper 10865.

Bayer, P. and C. Timmins. 2005. On the eqiulibrium properties of locational sorting models. *Journal of Urban Economics* 57: 462–477.

Bayer, P. and C. Timmins. 2007. Estimating equilibrium models of sorting. *The Economic Journal* 117: 353–374.

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

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

Berry, S. T. and P. A. Haile. 2010. Nonparametric identification of multinomial choice demand models with heterogeneous consumers. Working paper.

Berry, S. T. and P. A. Haile. 2014. Identification in differentiated products markets using market level data. *Econometrica* 82: 1749–1798.

Berry, S. T. and P. A. Haile. 2016. Identification in differentiated products markets. *Annual Review of Economics* 8:27–52.

Comscore. 2007, 2008, 2017, 2018. "Web Behavior Disaggregated: Dataset." Wharton Research Data Services, University of Pennsylvania.

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

Guiteras, R. P., J. Levinsohn, A. M. Mobarak. 2019. Demand estimation with strategic complementarities: sanitation in Bangladesh. BREAD working paper 553.

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

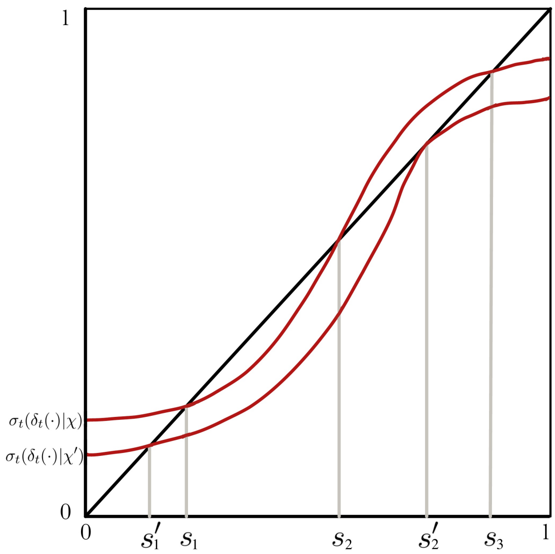
Rysman, M. 2004. Competition between networks: a study of the market for Yellow Pages. *Review of Economic Studies* 71: 483–512.

Timmins, C. and J. Murdock. 2007. A revealed preference approach to the measurement of congestion in travel cost models. *Journal of Environmental Economics and Management* 53: 230–249.

A Discussion of locally unique equilibria

This appendix discusses the concept of local uniqueness of equilibria as proposed in Subsection 3.4. In particular, I now build intuition for the condition that (5) must be nonsingular to ensure local uniqueness using a binary logit model with positive network externalities. Figure 7 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, roughly speaking, the equilibria s_2 and s_3 collapse into a single equilibrium s_2'. The matrix (5) 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 7: Multiple equilibria in a binary choice model with network externalities



B Identification appendix

B.1 Notes on the identification concept and proof techniques

This appendix fills in several details omitted in Section 4. It begins by discussing the role of Assumption ACDIST in the identification analysis of the market-data setting. It then discusses the NPIV framework that I use in my identification analysis and the order condition that I impose as a necessary condition for identification. It concludes by considering a generalization of the model considered in the identification analysis of the microdata setting in Section 4.

In its discussion of identification in the market-data setting, Section 4 includes Assumption ACDIST to ensure the invertibility of demand in the sense of Berry et al. (2013). Here, the invertibility of demand means that for all $\bar{\chi}$ in the support of χ, the function $\sigma(\cdot|\bar{\chi})$ is injective. Berry et al. provide the *connected substitutes* condition as a sufficient condition for the invertibility of demand. In this appendix, I show that Assumption ACDIST implies the connected substitues condition in my setting. Thus, when Assumption ACDIST holds, there exists a function $\sigma^{-1}(\cdot|\bar{\chi})$ defined on the image of $\sigma(\cdot|\bar{\chi})$ such that

$$\sigma^{-1}\left(\sigma(\delta|x) \mid \chi\right) = \delta. \tag{16}$$

Therefore, we can write

$$\sigma_j^{-1}(s|\chi) = g_j(s, x_j) + \xi_j \tag{17}$$

for each j.

The following lemma formally establishes Assumption ACDIST as a sufficient condition for the invertibility of demand.

Lemma 1. *Suppose Assumption ACDIST holds. Then, for all $\bar{\chi}$ in the support of χ, the function $\sigma(\cdot|\bar{\chi}) : \mathbb{R}^J \to \Delta^J$ is injective.*

Proof. In this proof, I suppress conditioning on χ in the market share function and write the market share function as $\sigma : \mathbb{R}^J \to \Delta^J$. Berry et al. (2013) provide three sufficient conditions for the invertibility of demand. The first is that the domain of σ is a Cartesian product, which holds in my setting. They label the second sufficient condition *weak substitutes*; it requires that $\sigma_j(\delta)$ is weakly decreasing in δ_k for all $j \in \{0, \ldots, J\}$ and $k \notin \{0, j\}$. This condition clearly holds in my model. Their third and final sufficient condition is the *connected strict substitution* condition, which they prove is equivalent to the following condition (adapted to my setting): for all $\delta \in \mathbb{R}^J$ and $\mathcal{K} \subseteq \{1, \ldots, J\}$, there exist $k \in \mathcal{K}$ and $\ell \notin \mathcal{K}$ such that $\sigma_\ell(\delta)$ is strictly decreasing in δ_k This condition holds in my setting because $\sigma_\ell(\delta)$ is strictly decreasing in δ_k whenever $k \neq \ell$ due to Assumption ACDIST. Therefore, σ is injective. □

The equations in (17) are the type of nonparametric regression equations studied by Newey and Powell (2003). Identification of model primitives via IVs in this setting requires the satisfaction of a completeness condition. In particular, consider the problem of identifying the function ϕ in

$$y = \phi(x, z_1) + \xi,$$

where (y, x, z_1) are observable random variables and ξ is an unobservable random variable. In this setting, identification in the IV framework requires variables z that satisfy $\mathbb{E}[\xi|z] = 0$ in addition to the following completeness condition: for all $\Gamma(x, z_1)$ with finite expectation, $\mathbb{E}[\Gamma(x, z_1)|z] = 0$ implies $\Gamma(x, z_1) = 0$. Now suppose that x_2 but not x_1 is included in z and write $z = (z_1, z_2)$, where z_2 is a d_{ex}-vector of excluded instruments. Also, let d_{en} be the dimension of x. Newey and Powell (2003) show that, in a model in which x is normally distributed conditional on z, a necessary order condition for the completeness condition is $d_{ex} \geq d_{en}$. They conjecture that such an order condition holds more generally. In this paper, I consider the function f to be identified by a nonparametric IV (NPIV) argument if (i) the existence of excluded instruments z_2 with $d_{ex} \geq d_{en}$ that satisfy $\mathbb{E}[\xi|z] = 0$ is possible under the model at hand and (ii) (x, z) satisfy the completeness condition.

In Section 4's discussion of identification with microdata, I consider indirect utilities of the form

$$u_{ijt} = g_j\left(w_{ijt},\, s_t,\, x_{jt}\right) + \xi_{jt} + \varepsilon_{ijt}, \tag{18}$$

for a scalar consumer characteristic w_{ijt}. In fact, this specification can be generalized without substantially changing the microdata identification results stated in Section 4. Indeed, consider indirect utilities of the form

$$u_{ijt} = g_j\left(w_{ijt}^{(1)},\, w_{ijt}^{(2)},\, s_t,\, x_{jt}\right) + \xi_{jt}\left(w_{ijt}^{(2)}\right) + \varepsilon_{ijt}, \tag{19}$$

where w_{ijt} are individual-level observables. In particular, $w_{ijt}^{(1)}$ is a scalar-valued individual characteristic and $w_{ijt}^{(2)}$ is a potentially multidimensional individual characteristic. This model captures heterogeneity in preferences for product characteristics across individuals in several ways. First, observable consumer characteristics w_{ijt} and product characteristics x_{jt} are permitted to interact in the nonparametric g_j functions. Second, the consumer's unobservable taste for good j in market t is permitted to vary according to his individual characteristics $w_{ijt}^{(2)}$. Third, the distribution of ε_{ijt} is permitted to depend on x_{jt}.

Following Berry and Haile (2010), we can relate the indirect utilities in (19) to those in (18) by conditioning on $w_{ijt}^{(2)}$, suppressing $w_{ijt}^{(2)}$ in our notation, and denoting $w_{ijt}^{(1)}$ by w_{ijt} in an abuse of notation. Then, the identification results from section 4 apply to the generalized model (19) as long as the conditions for those results are revised to include conditioning on $w_{ijt}^{(2)}$ as appropriate. Note that Assumption EPS-Micro rules out dependence between $w_{ijt}^{(1)}$ and ε_{ijt} but not between $w_{ijt}^{(2)}$ and ε_{ijt}, which allows for a great deal of flexibility in preference heterogeneity.

B.2 Identification with market data using restrictions on network externalities

Proposition 1 establishes that the assumption of a known distribution for the idiosyncratic part of consumers' utilities, i.e. Assumption KD, enables the identification of the model with market data alone. An alternative assumption (which is similar to Assumption SREG-Market) also allows for the identification of the semi-nonparametric model with market data. This assumption, which I now state, places a significant restriction on the form of network externalities.

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

$$g_j(s, x_j) = \alpha s_j + \tilde{g}_j(x_j),$$

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.

Proposition 7. *Suppose that Assumption SEPSHARE holds. Suppose additionally that $J > 1$ and that the dimension of x_j, call it d_x, is constant across j and strictly exceeds one. Then, g, σ, and ξ are identified.*

Proof. I first consider identification in case (i), in which

$$\sigma_j^{-1}(s) = s_j + \tilde{g}_j(x_j) + \xi_j$$

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

$$s_j = \underbrace{\sigma_j^{-1}(s) - \tilde{g}_j(x_j)}_{=:\kappa_j(s,x_j)} + \xi_j. \tag{20}$$

Equation (20) falls within the NPIV framework: the left-hand side is observable and $\kappa_j(s, x_j)$ is an unknown nonparametric function of (s, x_j). Consequently, the function κ_j is identified by an NPIV argument if there are instruments z_j that satisfy the mean-independence assumption $\mathbb{E}[\xi_j|z_j] = 0$ and following completeness condition: for all functions $\Gamma : \Delta^J \times \operatorname{supp} x_j \to \mathbb{R}$ such that $\mathbb{E}|\Gamma(s, x_j)| < \infty$, $\mathbb{E}[\Gamma(s, x_j)|z_j] = 0$ (almost surely) implies $\Gamma(s, x_j) = 0$ (almost surely). Recall that I require the following order condition to apply an NPIV argument for identification: the dimension of the instruments z_j must be at least as great as the dimension of (s_t, x_j), i.e. of at least dimension $J + d_x$. The instruments $z_j = (x_1, \ldots, x_J)$ satisfy the mean independence condition by my assumption in (11). Furthermore, the dimension of z_j is $J \cdot d_x$, which is greater than $J + d_x$ since we have assumed that $J > 1$ and $d_x > 1$. Therefore, κ_j is identified. The identification of ξ_j follows immediately. For the separate identification of σ_j^{-1} and $\tilde{g}_j(x_j)$, 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 $\tilde{g}_j(\bar{x}_j) = 0$. Then, $\sigma_j^{-1}(s) = \kappa_j(s, \bar{x}_j)$ and $\tilde{g}_j(x_j) = \kappa_j(s, x_j) - \sigma_j^{-1}(s)$. Next, note that σ_j^{-1} is differentiable as a consequence of Assumption ACDIST and that all partial derivatives of σ_j^{-1} are identified by the separability of σ_j^{-1} and $\tilde{g}_j$ in the definition of κ_j. Pick some $s^\dagger$ in the support of $s|\bar{x}_j$. Then, $\sigma_j^{-1}(s^\dagger) = \kappa_j(s^\dagger, \bar{x}_j)$. Therefore,

$$\begin{aligned}\int_{s^\dagger}^{s} \frac{\partial \sigma_j^{-1}(\tilde{s})}{\partial s} d\tilde{s} + \kappa_j(s^\dagger, \bar{x}_j) &= \sigma_j^{-1}(s) - \sigma_j^{-1}(s^\dagger) + \kappa_j(s^\dagger, \bar{x}_j) \\ &= \sigma_j^{-1}(s),\end{aligned}$$

which identifies σ_j^{-1} on the entire support of s and consequently identifies $\tilde{g}_j(x_j) = \kappa_j(s, x_j) - \sigma_j^{-1}(s)$ on the full support of x_j. □

The result above is significant because it shows that a basic model of network externalities is identified by a substative 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 requires the econometrician to impose 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.

B.3 Proof of Proposition 1

Proof. We can invert choice probabilities as to obtain

$$\sigma_j^{-1}(s) = g_j(s, x_j) + \xi_j \tag{21}$$

for all j, where σ_j^{-1} is known.

Identification under completeness requires the availability of J excluded instruments for the J-dimensional endogenous regressor s. Given that $\{x_j\}_{j \in \mathcal{J}}$ are the only observable variables that determine s, the only candidates for excluded instruments are x_{-j}, of which there are $J - 1$. Therefore, the order condition stated in Appendix B.1 fails and consequently g_j and ξ_j are not identified without additional restrictions.

Assumption EXCL provides an additional restriction that enables the identification of g_j and ξ_j. Assumption EXCL allows us to write (in an abuse of notation) $g_j(s, x_j) = g_j(s_{-\ell}, x_j)$, where ℓ is the good whose share is excluded from g_j. This reduces the number of endogenous regressors to $J-1$, which is the number of available excluded regressors. Therefore, the order condition required for the NPIV argument is satisfied and $g_j(s, x_j)$ and ξ_j are consequently identified. □

Identification here follows from a standard NPIV argument. The exclusion restriction enables identification by providing an excluded instrument x_ℓ that affects good j's market share only through a competitive effect (i.e. by changing good ℓ's mean value to consumers) rather than through network externalities. This allows for the separation of these effects and thus the identication of network externalities.

B.4 Proof of Proposition 2

Proof. Let

$$\rho_{ij0} = \Pr\left(u_{ijt} \leq 0 \quad \forall j \in \mathcal{J}_t \setminus \{0\} \mid t,\, w_{it}\right).$$

Conditioning on t here involves conditioning on (x_t, s_t, ξ_t) jointly. Note that ρ_{ij0} is the probability of a member of market t with observable characteristics w_{it} choosing the outside option; this quantity is known from the joint distribution of observables. Next, define

$$\begin{aligned}\tilde{u}_{ijt} &= u_{ijt} - w_{ijt} \\ &= \delta_{jt} + \tilde{\varepsilon}_{ijt}\end{aligned}$$

and note that

$$\begin{aligned}\rho_{ij0} &= \Pr\left(\delta_{jt} + \tilde{\varepsilon}_{ijt} \leq -w_{ijt} \quad \forall j \in \mathcal{J}_t \setminus \{0\} \mid t,\, w_{it}\right) \\ &= F_{\tilde{u}|t}(-w_{it}),\end{aligned}$$

where $F_{\tilde{u}|t}$ is the distribution function of $\tilde{u}_{it}$ (which is an $\mathbb{R}^J$-valued random vector) in market t. Given the large-support assumption in the prompt, the above establishes the identification of the distribution of $\tilde{u}_{ijt}$ in each market t. Since I made the location normalization that $\mathbb{E}[\tilde{\varepsilon}_{ij}|t] = 0$, the mean of $\tilde{u}_{ijt}$ conditional on t is δ_{jt}, which establishes the identification of δ_{jt}. Identification of the distribution of the $\tilde{u}_{ijt}$ together with identification of the δ_{jt} implies the identification of $G_{\varepsilon|x}$. □

B.5 Proof of Proposition 3

As shown in Appendix B.3, identification in the setting where the econometrician knows the distribution of the ε_{ijt} deviates followed from a standard nonparametric IV argument with $\sigma_j^{-1}(s)$ as the observed dependent variable and endogenous regressors s. With σ_j^{-1} unknown, this argument does not apply; in fact, as Proposition 3 states, the model primitives are generally not identified without additional assumptions.

Proof. Let σ denote the market share function under F_ε and let $\tilde{\sigma}$ denote the market share function under $\tilde{F}_\varepsilon$. For $(F_\varepsilon,\, g)$ to be consistent with the model and observable data, we need

$$\sigma_j^{-1}(s) = g_j(s, x_j) + \xi_j, \quad j \in \mathcal{J}. \tag{22}$$

Consider

$$\tilde{g}_j(s, x_j) = g_j(s, x_j) + \tilde{\sigma}_j^{-1}(s) - \sigma_j^{-1}(s). \tag{23}$$

Solving (23) for g_j, substituting g_j into (22), and re-arranging terms yields

$$\tilde{\sigma}^{-1}(s) = \tilde{g}_j(s, x_j) + \xi_j, \quad j \in \mathcal{J},$$

which implies that $\tilde{\theta} = (\tilde{F}_\varepsilon,\, \tilde{g})$ supports the same market shares as θ for a given (x, ξ). This confirms that $\tilde{\theta} = (\tilde{F}_\varepsilon,\, \tilde{g})$ is observationally equivalent to θ. Now note that if the functions $\{g_j\}$ satisfy Assumption SREG-Market, then so will the $\{\tilde{g}_j\}$ functions. This establishes that the result remains true under Assumption SREG-Market. □

B.6 Proof of Proposition 6

Proof. I first consider identification in case (i). The proof is almost identical to that of Proposition 7, except that M_t is available as an additional instrument when the assumption in (11) holds.

Identification under case (ii) follows a similar argument. In this case, we have it that

$$\sigma_j^{-1}(s) = x_j + \tilde{g}_j(Rs) + \xi_j$$

and thus

$$x_{jt} = \sigma_j^{-1}(s_t) - \tilde{g}_j(M_t s_t) - \xi_{jt}.$$

Identification of

$$\kappa_j(s_t, M_t) = \sigma_j^{-1}(s_t) - \tilde{g}_j(M_t s_t)$$

under completeness follows from (20) with the instruments $z_{jt} = (x_{1t}, \dots, x_{J_t t}, M_t)$, of which there are $J_t + 1$. This is the same number of regressors in (s, M_t).

I first consider separate identification of $\tilde{g}$ and σ under assumption (a), the "own shares" assumption. Here, we note that

$$\begin{aligned}
\frac{\partial \kappa_j}{\partial s_k}(s, M) &= \frac{\partial \sigma_j^{-1}}{\partial s_k}(s), \qquad k \neq j \\
\frac{\partial \kappa_j}{\partial s_j}(s, M) &= \frac{\partial \sigma_j^{-1}}{\partial s_j}(s) - \tilde{g}_{jj}(Ms)M \\
\frac{\partial \kappa_j}{\partial M}(s, M) &= -\tilde{g}_{jj}(Ms)s_j
\end{aligned}$$

where $\tilde{g}_{jk}$ denotes the derivative of $\tilde{g}_j$ with respect to its kth argument. Since we have identified κ_j, the equations above imply the identification of $\tilde{g}_{jj}$ and all of the partial derivatives of σ_j^{-1}. To separately identify the levels of these functions, we require a location normalization, e.g. there is a $\bar{s}$ known to the researcher such that $\sigma_j^{-1}(\bar{s}) = 0$. Then, we can integrate the identified partial derivatives to obtain σ_j^{-1}; the identification of $\tilde{g}_j$ then follows.

I now consider separate identification of $\tilde{g}$ and σ under assumption (b), the "exchangeability and exclusion" assumption. Differentiating κ_j with respect to M yields

$$\frac{\partial \kappa_j}{\partial M}(s, M) = -\sum_{k=1}^{J} \tilde{g}_{jk}(Ms)s_j. \tag{24}$$

where $\tilde{g}_{jk}$ denotes the derivative of $\tilde{g}_j$ with respect to its kth argument. Recall that $\tilde{g}_{j\ell} = 0$ by assumption. Thus, differentiating κ_j with respect to s_ℓ yields

$$\frac{\partial \kappa_j}{\partial s_\ell}(s, M) = \frac{\partial \sigma_j^{-1}}{\partial s_\ell}(s).$$

Therefore, $\partial \sigma_j^{-1}(s)/\partial s_\ell$ is identified. By my exchangeability assumption, this identifies all cross-partials of σ_j^{-1}. Therefore, we can identify the cross-partials of $\tilde{g}_j$ via

$$\frac{\partial \kappa_j}{\partial s_r}(s, M) = \frac{\partial \sigma_j^{-1}}{\partial s_r}(s) - \tilde{g}_{jr}(Ms, x_j)$$

for $r \in \mathcal{J} \setminus \{j\}$. With each of the $\tilde{g}_{jr}$ identified, (24) pins down $\tilde{g}_{jj}(Ms, x_j)$. Then,

$$\frac{\partial \kappa_j}{\partial s_j}(s, M) = \frac{\partial \sigma_j^{-1}}{\partial s_j}(s) - \tilde{g}_{jj}(Ms)M$$

identifies $\partial\sigma_j^{-1}(s|x)/s_j$. Given that we have identified all partial derivatives of σ_j^{-1}, we can identify its level after a location normalization. □

B.7 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/product characteristics that vary within demograpic 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 $J_t \times D$ matrix whose dth column provides the market shares of the J_t inside goods among members of demographic group d in market t. I do not consider the S-type model in the submarket data case, as it is plagued by the same identification problems as the S-type model without multiple demographic groups in the market data setting. To see why, note that the the S-type model with distinct demographic groups collapses to the S-type model without multiple demographic groups in the special case in which consumers only care about market shares within their own demographic groups and we take the "markets" to be pairs of markets and demographic groups. Therefore, I begin by considering the Q-type model in the submarket data setting. This model features indirect utilities of the form

$$u_{ijt} = g_j^{d(i)}(S_t M_t^\dagger, x_{jt}) + \xi_{jt}^{d(i)} + \varepsilon_{ijt}$$

where $d(i)$ denotes consumer i's demographic group and

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

so that

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

This model permits both the market/choice unobservables ξ_{jt}^d and the structural function g_j^d to vary across demographic groups. In addition, I do not impose that the distribution of ε_{ijt} is constant across demographic groups. Under conditions analogous to those stated in previous subsections, we can invert demographic-group specific market shares to obtain

$$\sigma_{j,d}^{-1}(s_t^d) = g_j^d(S_t M_t^\dagger, x_{jt}) + \xi_{jt}^d + \varepsilon_{ijt}.$$

Note that we have not yet imposed a scale normalization, which will be necessary for identification. We now express x_{jt} as $x_{jt} = \{x_{jt}^d\}_{d \in D}$ and assume that

$$g_j^d(S_t, x_{jt}) = x_{jt}^d + \tilde{g}_j^d(S_t M_t^\dagger).$$

for some functions $\tilde{g}_j^d$. The 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. This is because it remains possible to set $x_{jt}^d = x_{jt}$ for the product-market characteristic x_{jt}. The assignment of a coefficient of one to x_{jt}^d is not a substantive assumption but instead a scale normalization.

We can now write

$$x_{jt}^d = \underbrace{\sigma_{j,d}^{-1}(s_t^d) - \tilde{g}_j^d(S_t M_t^\dagger)}_{=:\kappa_j^d(S_t, M_t)} + \xi_{jt}^d.$$

This is an NPIV-style regression equation with $J_t D + D$ regressors. Possible instruments include the D-dimensional M_t and the BLP instruments $\{x_{jt}^d : d \in D, j \in \mathcal{J}_t\}$, of which there are $J_t D$. Note that for the order condition stated in Appendix B.1 to hold, the x_{jt}^d must be distinct across d and j; that is, we need shifters for specific pairs of products and demographic groups. Although we can reduce our instrument requirements by assuming that $\tilde{g}_j^d$ 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 $s_t^{d'}$.

As long as a completeness condition and an exclusion restriction hold, we can argue that κ_j^d and ξ_{jt}^d are identified. We can then use, as in Proposition 6, a combination of exclusion restrictions and exchangeability assumptions to separately identify $\sigma_{j,d}^{-1}$ and $\tilde{g}_j^d$. This separate identification critically relies on the fact that overall demand rather than market shares enter consumers' indirect utilities.

I now discuss identification in the microdata setting. Identification analysis in this setting is very similar to identification analysis in the microdata setting without distinct demographic groups. Although I study a Q-type model in what follows, the identification discussion that I present naturally extends to an S-type model. In the microdata setting, I assume that consumers have characteristics $\{w_{ijt}\}_{j \in \mathcal{J}_t}$ that vary within demographic groups. In particular, I assume that w_{ijt} has a large support within each demographic group and write consumers' indirect utilities as

$$u_{ijt} = w_{ijt} + \tilde{g}_j^{d(i)}(S_t M_t^\dagger, x_{jt}) + \xi_{jt}^{d(i)} + \varepsilon_{ijt}.$$

In addition, I define (suppressing market subscripts)

$$\delta_j^d = \tilde{g}_j^d(SM^\dagger x_j) + \xi_j^d. \tag{25}$$

A standard nonparametric discrete-choice argument as in the proof of Proposition 4 identifies δ_j^d and the distribution of the ε_{ijt}. We then analyse (25) in the NPIV framework. The regressors are x_j, which I assume has constant dimension d_x across j, and $SM^\dagger$, which has dimension JD. Possible instruments include (i) the product/market characteristics $\{x_j\}_{j \in \mathcal{J}}$, (ii) the market size vector M; (iii) moments of the distribution of w_{ij} within each demographic group. Note that these are essentially the same instruments as are available in the case without distinct demographic groups. We achieve identification by assuming that these instruments satisfy the requisite exclusion and completeness conditions, which are generalizations of Assumptions NPIV-EX and NPIV-C to the $D > 1$ case.

C Price sensivity estimation

This section provides additional details of the price-sensitivity estimation procedure described in Subsection 5.3 To begin, I define $\bar{\sigma}_{jt}(p,s)$ as site j's market share in market t at prices p and market shares s. This is simply the mean probability of choosing good j under prices p and market shares s, where the mean is taken over consumers in market t (i.e. 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 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,

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

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 for a particular site j:

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

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 anaolgoues 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

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

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 first-order condition (15) 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 S-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 first-order condition 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 $D_p\sigma_{jt}(p;\alpha)$:

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

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 prive derivatives are $JD \times J$ and (ii) the specification above includes both S-type and Q-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 estiamtion.

D Additional results

This appendix provides additional results from the paper's empirical analysis. First, Table 19 provides results of the microstep estimation for the "Race" demographic group specification. I did not include the race regressors in this regression. Next, Table 20 provides results of the microstep estimation for the "Age" demographic group specification.

Table 19: First-stage parameter estimates ("Race" demographic group specification)

	eharmony.com	match.com	okcupid.com	pof.com	zoosk.com
Education: High school or less (Omit.)	0.000	0.000	0.000	0.000	0.000
	-	-	-	-	-
Education: Some college	-0.113	0.011	0.087	0.119	-0.257
	(0.119)	(0.072)	(0.281)	(0.184)	(0.514)
Education: College degree	-0.099	-0.004	0.329	0.096	-0.252
	(0.123)	(0.074)	(0.267)	(0.185)	(0.491)
Education: Advanced degree	0.062	0.081	-0.363	-0.248	-0.603
	(0.140)	(0.086)	(0.415)	(0.266)	(1.105)
Education: Unknown	0.052	-0.178	-0.322	0.208	-1.428
	(0.086)	(0.054)	(0.224)	(0.145)	(0.550)
Age: Under 25yo (Omit.)	0.000	0.000	0.000	0.000	0.000
	-	-	-	-	-
Age: 25-29yo	0.015	0.069	-0.018	0.208	3.731
	(0.141)	(0.096)	(0.361)	(0.216)	(6.396)
Age: 30-34yo	-0.087	0.071	-0.331	0.363	4.998
	(0.133)	(0.089)	(0.349)	(0.199)	(6.333)
Age: 35-39yo	0.025	0.036	-0.269	0.198	4.611
	(0.129)	(0.088)	(0.334)	(0.199)	(6.338)
Age: 40-49yo	0.008	-0.002	-0.137	-0.024	4.196
	(0.122)	(0.083)	(0.308)	(0.189)	(6.328)
Age: 50+yo	-0.036	0.001	-0.285	0.078	4.740
	(0.121)	(0.082)	(0.305)	(0.186)	(6.322)
Children in HH: No (Omit.)	0.000	0.000	0.000	0.000	0.000
	-	-	-	-	-
Children in HH: Yes	0.024	0.001	0.131	-0.053	-0.076
	(0.052)	(0.035)	(0.149)	(0.082)	(0.365)
Broadband: No (Omit.)	0.000	0.000	0.000	0.000	0.000
	-	-	-	-	-
Broadband: Yes	0.532	1.093	1.220	0.247	4.451
	(0.081)	(0.068)	(0.339)	(0.118)	(9.298)
Hispanic: No (Omit.)	0.000	0.000	0.000	0.000	0.000
	-	-	-	-	-
Hispanic: Yes	-0.369	-0.281	-0.560	-0.457	-0.480
	(0.045)	(0.029)	(0.136)	(0.075)	(0.338)
Income: Under 25k (Omit.)	0.000	0.000	0.000	0.000	0.000
	-	-	-	-	-
Income: 25-75k	-0.084	-0.106	-0.103	-0.001	-0.014
	(0.050)	(0.034)	(0.146)	(0.078)	(0.355)
Income: 75-100k	-0.104	-0.144	-0.115	-0.127	-0.458
	(0.062)	(0.042)	(0.180)	(0.102)	(0.541)
Income: Over 100k	-0.230	-0.158	-0.088	-0.395	-0.427
	(0.061)	(0.040)	(0.167)	(0.101)	(0.440)
HH size: 1 (Omit.)	0.000	0.000	0.000	0.000	0.000
	-	-	-	-	-
HH size: 2	-0.018	-0.002	-0.172	-0.008	-0.253
	(0.076)	(0.051)	(0.200)	(0.114)	(0.400)
HH size: 3	-0.155	-0.077	-0.306	0.008	-0.285
	(0.091)	(0.061)	(0.244)	(0.137)	(0.501)
HH size: Over 3	-0.050	-0.093	-0.294	-0.053	-0.571
	(0.089)	(0.060)	(0.236)	(0.134)	(0.491)

Table 20: First-stage parameter estimates ("Age" demographic group specification)

	eharmony.com	match.com	okcupid.com	pof.com	zoosk.com
Education: High school or less (Omit.)	0.000	0.000	0.000	0.000	0.000
	-	-	-	-	-
Education: Some college	-0.118	0.009	0.091	0.096	-0.229
	(0.119)	(0.072)	(0.282)	(0.185)	(0.514)
Education: College degree	-0.110	-0.005	0.336	0.051	-0.208
	(0.123)	(0.074)	(0.268)	(0.186)	(0.492)
Education: Advanced degree	0.050	0.073	-0.330	-0.270	-0.564
	(0.140)	(0.086)	(0.415)	(0.266)	(1.105)
Education: Unknown	0.047	-0.183	-0.316	0.185	-1.430
	(0.086)	(0.055)	(0.224)	(0.145)	(0.552)
Age: Under 25yo (Omit.)	0.000	0.000	0.000	0.000	0.000
	-	-	-	-	-
Age: 25-29yo	0.009	0.058	-0.058	0.146	4.751
	(0.142)	(0.096)	(0.365)	(0.218)	(9.720)
Age: 30-34yo	-0.099	0.050	-0.389	0.295	6.036
	(0.133)	(0.090)	(0.353)	(0.202)	(9.678)
Age: 35-39yo	-3.169	-2.672	-5.502	-3.801	-48.014
	(0.135)	(0.096)	(0.432)	(0.210)	(4.287)
Age: 40-49yo	-3.183	-2.708	-5.385	-4.026	-48.419
	(0.129)	(0.093)	(0.415)	(0.202)	(4.278)
Age: 50+yo	-3.233	-2.709	-5.532	-3.930	-47.861
	(0.127)	(0.092)	(0.412)	(0.198)	(4.272)
Children in HH: No (Omit.)	0.000	0.000	0.000	0.000	0.000
	-	-	-	-	-
Children in HH: Yes	0.024	0.003	0.129	-0.040	-0.086
	(0.053)	(0.035)	(0.149)	(0.083)	(0.364)
Race: White (Omit.)	0.000	0.000	0.000	0.000	0.000
	-	-	-	-	-
Race: Black	0.003	-0.491	-0.611	-0.418	-0.189
	(0.066)	(0.053)	(0.233)	(0.119)	(0.427)
Race: Asian	-0.132	-0.332	-0.162	-0.775	0.297
	(0.149)	(0.096)	(0.297)	(0.259)	(0.397)
Race: Other	-0.230	-0.315	0.063	-0.249	0.239
	(0.142)	(0.088)	(0.258)	(0.168)	(0.389)
Broadband: No (Omit.)	0.000	0.000	0.000	0.000	0.000
	-	-	-	-	-
Broadband: Yes	0.535	1.100	1.225	0.245	4.870
	(0.081)	(0.068)	(0.341)	(0.119)	(10.665)
Hispanic: No (Omit.)	0.000	0.000	0.000	0.000	0.000
	-	-	-	-	-
Hispanic: Yes	-0.368	-0.277	-0.561	-0.457	-0.521
	(0.046)	(0.030)	(0.136)	(0.075)	(0.357)
Income: Under 25k (Omit.)	0.000	0.000	0.000	0.000	0.000
	-	-	-	-	-
Income: 25-75k	-0.098	-0.117	-0.100	-0.012	-0.025
	(0.050)	(0.034)	(0.147)	(0.079)	(0.355)
Income: 75-100k	-0.120	-0.161	-0.120	-0.141	-0.477
	(0.063)	(0.043)	(0.182)	(0.102)	(0.542)
Income: Over 100k	-0.245	-0.175	-0.091	-0.409	-0.448
	(0.061)	(0.040)	(0.169)	(0.102)	(0.440)
HH size: 1 (Omit.)	0.000	0.000	0.000	0.000	0.000
	-	-	-	-	-
HH size: 2	-0.024	-0.008	-0.170	-0.015	-0.240
	(0.076)	(0.051)	(0.200)	(0.114)	(0.400)
HH size: 3	-0.159	-0.082	-0.303	-0.003	-0.282
	(0.091)	(0.061)	(0.244)	(0.137)	(0.500)
HH size: Over 3	-0.056	-0.099	-0.295	-0.066	-0.562
	(0.089)	(0.060)	(0.236)	(0.134)	(0.490)