Demand with network externalities: identification and an application to the dating websites industry*

Michael Sullivan

Department of Economics, Yale University

November 1, 2022

Abstract

This paper characterizes the identifiability of demand models with network externalities. Guided by my identification analysis, I empirically evaluate how network externalities shape the effects of consolidation in the US dating websites industry. Network externalities often arise in differentiated products markets, and especially in platform markets. I show that demand models with network externalities are generally not identified with market-level data alone. This result reflects the impossibility of independently varying product characteristics and market shares at the market level. However, straightforward extension of results in Berry and Haile (2022) establishes that demand models with network externalities are identified under reasonable conditions with microdata linking consumers' decisions and characteristics. I estimate demand for dating websites using online browsing microdata. Under my preferred estimates, a user of a site values a 10% increase in the site's usership at \$6.34/month. I find that welfare losses from increased prices outweigh the gains from network externalities associated with monopolization. Additionally, I find that a firm earns higher profits from joint ownership of the two largest dating websites when it does not integrate these sites.

*Email address: `m.r.sullivan@yale.edu`. I thank Phil Haile, Steven Berry, Katja Seim, Mitsuru Igami, and Judy Chevalier as well as participants in Yale's Industrial Organization Prospectus Workshop for helpful feedback on this project. I additionally thank the Tobin Center for Economic Policy at Yale University and the Yale School of Management for their support in acquiring data used in this project. This paper draws on research supported by the Social Sciences and Humanities Research Council.

1 Introduction

This paper characterizes the identifiability of demand models with network externalities. Guided by my identification analysis, I empirically evaluate how network externalities shape the welfare effects of consolidation in the US dating websites industry. Network externalities, which arise when a product's value to a consumer depends on other consumers' usage of the product, often affect demand for differentiated products. They are especially relevant in platform markets wherein the selection of agents available for interaction on a platform underlies the platform's value. When network externalities affect consumer demand, they change standard analyses of competition. Consumers who enjoy using the same social media service as other consumers, for example, may benefit from the consolidation of social media services despite the capacity for consolidation to induce quality reduction or price hikes.

The analysis of demand faces identification challenges in the presence of network externalities. These challenges reflect the fact that a change in a product characteristic both directly and indirectly affects consumer choices. First, consumers may intrinsically value the product characteristic; a product's price reduction, for example, may draw consumers to use the product (all else equal). The change in product usage owing to this direct effect of changing a product characteristic may further shift the product's appeal to consumers on account of network externalities; the new consumers that a price reduction draws to a product, for example, make the product more appealing to other consumers under positive network externalities. Disentangling the direct effect and network externalities is often impossible with only market-level data on product characteristics and sales; this reflects the impossibility of varying product characteristics while holding market shares fixed to identify product characteristics' effects on consumer choices.

The availability of microdata linking individual consumers' choices with their characteristics, however, improves the prospects for identification. Within- and cross-market variation in consumer characteristics provide different information about (i) direct effects of consumer characteristics on consumer choice probabilities and (ii) network externalities. The first of these effects are straightforwardly identified by varying consumer characteristics within a market. Consumers with fast internet connections, for example, may be more likely to use a dating website than consumers in the same city with slow connections. Varying a market's distribution of consumer characteristics, meanwhile, has two effects on product usage: the aforementioned direct effect identified from within-market variation, and an indirect effect reflecting network externalities. This indirect effect arises, for example, when an increase in the share of a city with fast internet leads more consumers to use dating websites, thereby making the dating website more appealing to other consumers. When the direct effect is identified from within-market variation, cross-market variation identifies the indirect effect under appropriate assumptions on the nature of variation across markets. My approach uses instrumental variables that capture this variation. The first set of available instruments includes functions of market-specific distributions of individual characteristics, e.g., the share of consumers in a city with high-speed internet. The validity of these instruments requires that their associated distributions of consumer characteristics to not directly affect consumer tastes. This condition is violated when firms set local product quality or advertising in markets based on local consumer characteristics and when consumers' tastes directly depend on the characteristics of their neighbours. The second set of instruments includes characteristics of other products in the market, i.e., the instruments of Berry et al. (1995). These are generally the only two types of instruments that are consistent with the demand model considered by this paper.

My identification findings guide my empirical analysis of the US dating websites industry in 2007–2008. The industry witnessed considerable consolidation in the decade following this time period, with market leader Match acquiring competitors OKCupid in 2011 and Plenty of Fish in 2015. Also, Match competitor Spark Neworks acquired the dating site Zoosk in 2019. These acquisitions did not immediately lead to the integrations of the acquiring firms' platforms with the acquired platforms. This raises the question of whether network externalities could make acquisitions followed by platform integrations welfare enhancing, and of whether pronounced platform differentiation could explain firms' hesitancy to integrate platforms following acquisitions. An additional feature of the dating websites industry is that market shares exhibit considerable variation across geographically defined markets; this is consistent with different markets tipping toward different sites on account of network externalities. Another possibility is that taste differences across markets induce cross-market variation in market shares.

I study demand for dating websites using online browsing microdata that provides consumer locations, characteristics, and records of browsing dating websites. Given my theoretical conclusion that within-market variation in consumer characteristics is important for identification, I specify a model of demand with rich consumer heterogeneity based on observable characteristics. Additionally, I estimate demand using instrumental variables. I use instruments based on market-specific distributions of consumer characteristics, and I choose the consumer characteristics on which I base my instrumental variables to mitigate concerns that these variables will violate the exclusion restriction required for their validity. These characteristics include an indicator for whether the consumer has broadband internet and measures of consumers' internet usage.

My estimates suggest that network externalities are substantial and account for most the cross-market variation in sites' market shares. Under the estimates of my preferred specification, an inframarginal user of a site values a 10% increase in the site's usership at \$6.34/month, which is about one third of the most popular site's price. Moreover, my estimates provide evidence of age-based homophily in dating website choice. I use my counterfactual model to decompose variation in dating website usage across geographical markets into network externalities, unobserved local taste differences, and differences in consumer characteristics; I find that most of this variation owes to network externalities. My estimated model also facilitates my assessment of the welfare effects of consolidation in the dating websites industry, i.e., whether network externalities are strong enough to counteract the usual harms to consumers associated with consolidation related to market power. Counterfactual analysis conducted with my estimated model suggests that welfare losses from increased prices exceed the gains from network externalities associated with a move to monopoly in the dating website industry. These results suggest that the importance of network externalities in platform markets does not nullify the usual importance of price responses in shaping the welfare effects of consolidation. Additionally, I find that a firm earns higher profits (excluding fixed costs) from joint ownership of the two largest US dating websites than from integrating these dating websites into a single site. This outcome reflects that a portfolio of differentiated website is more attractive to consumers than a single website with an especially large user base. The strength of platform differentiation relative to network externalities may explain for why dating website acquisitions have not been followed by website integrations in practice.

1.1 Related literature

The literature on network externalities distinguishes between direct and indirect network externalities. Direct network externalities emerge when an agent's pay-offs directly depend on the choices of other agents facing the same choice problem, whereas indirect network externalities arise when engagement with a product induces changes in the product's attractiveness. Indirect network externalities are typical of two-sided platforms; see Jullien et al. (2021) for discussion of these externalities. My study focuses on direct network externalities, which are typical of social media and communication platforms.

There is a sizeable literature on the identification of network externalities, broadly defined. Many such papers concern themselves with "peer effects"; some notable papers in this category include Manski (1993), Angrist (2014), Graham (2008), Graham (2018), and Bramoullé et al. (2009). These papers investigate the determination of an individual's (typically continuous) outcome by observable covariates, characteristics of the individual's peers, and an unobservable shifter of the outcome of interest. My paper considers a different setting in which individuals' choices among discrete alternatives are simultaneously determined in light of network externalities, alternative-specific characteristics, and alternative-specific unobservables. Additionally, my paper's static setting contrasts to the dynamic setting studied by, e.g., Kim et al. (2021). The theoretical literature that is most closely related to my work is that on discrete choice models of social interactions. This literature is reviewed by Durlauf and Ioannides (2010), and many of its studies have been authored by Brock and Durlauf. See, for example, Brock and Durlauf (2001a), Brock and Durlauf (2001b), and Brock and Durlauf (2007). This paper's analysis differs from this literature in several ways. First, I analyze a setting with a continuum of decision-makers, as is typically specified in discrete-choice demand models in industrial organization. Second, I compare two common settings in demand estimation in industrial organization: the setting in which the research has access to data on quantities and product characteristics at the market level ("market data"), and the setting in which the research has access to data on individual consumers' decisions and characteristics ("microdata"). My paper also focuses on models of differentiated products demand of the sort reviewed by **?**. I especially emphasize the role of these models' structural product/market-level unobservables, which are often denoted ξ_{jt} and which are the source of the econometric endogeneity problem in differentiated products demand models. My paper makes no restriction on the distribution of alternative/market-level unobservables; contrast this approach to Brock and Durlauf (2007), who focus on cases in which these unobservables are uniformly zero or are otherwise restricted.

My paper also relates to the empirical literature on consumer choice with network externalities. Some empirical papers that operate in a setting similar to that of my paper include Timmins and Murdock (2007) (which follows the model and estimation procedure of Bayer and Timmins (2007)), Bayer et al. (2004), Guiteras et al. (2019), and Allende (2019). My work is somewhat less related to empirical studies of two-sided markets with indirect network externalities, e.g., Rysman (2004) and Farronato et al. (2020).

My paper relates to several other literatures. First, I use techniques from the literature on the identification of demand for differentiated products, including Berry et al. (2013), Berry and Haile (2014), Berry and Haile (2016), and especially Berry and Haile (2022). My paper also relates to the literature on matching markets—e.g., Smith (2006)—which provides a microfoundation for network externalities in the dating websites industry. Last, my paper relates to the literature on online dating. This literature has mostly focused on activity within a dating website as opposed to the role of network externalities in driving choice of dating website, e.g., Hitsch et al. (2010) and Fong (2020).

2 Setting and data

This paper's empirical application studies the US dating websites industry in 2007–2008 using the Comscore Web Behavior Database. This dataset includes the browsing and online transactions records for a large panel of US households (91689 panelists in 2007 and 57817 panelists in 2008). 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. Most individuals in the Comscore data appear in only one year. When an individual appears in multiple years, I treat that individual's records for each year as a separate panelist. Note that De Los Santos et al. (2012) find that individuals in Web Behavior Database are representative of online buyers in the US.

For each 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 1 reports the most-used dating sites in 2007–2008. Each row of the "Share (%)" column

Table 1: Most popular dating websites sites, 2007–2008

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

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

reports the share of households in the Comscore data who spend at least five minutes on the indicated site. Most dating website usage in my sample is accounted for by a few major sites that appeal to broad audiences. These include `match.com`, `eharmony.com`, `pof.com` ("Plenty of Fish"), and `okcupid.com`, which are the sites that I focus on in my main analysis. A smaller but not insubstantial amount of usage is accounted for by sites appealing to various subpopulations (e.g. `catholicmatch.com` for Catholic users and `silversingles.com` for older users). Of the sites on which I focus in my empirical analysis, `eharmony.com` and `match.com` require payment for use whereas `pof.com` and `okcupid.com` are free to use. These free sites rely on advertising and paid premium features for revenue. The monthly prices of subscriptions for `eharmony.com` and `match.com` in 2007 were \$59.95 and \$34.99, respectively.

The markets in my analysis are geographical regions based on combined statistical areas (CSAs), which consist of counties that are economically and/or socially connected. I construct CSA/state pairs by dividing each CSA up into its parts that belong to different states. 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 and that does not belong to a CSA/state to this county's closest CSA/state.

I assign each panelist in my estimation sample to either a primary site or the outside option. A panelist who visits at least five pages within a dating website, visits the site during at least two distinct sessions, and spends at least five minutes on the site qualifies as a user of that site. A panelist's primary site is the site on which the panelist spends the most time

Table 2: Multihoming patterns

Website	Share of panelists using	Share also using			
		`eharmony.com`	`match.com`	`okcupid.com`	`pof.com`
eharmony	0.35	1.00	0.38	0.02	0.09
match	0.67	0.20	1.00	0.02	0.07
okcupid	0.04	0.19	0.26	1.00	0.19
pof	0.15	0.21	0.32	0.05	1.00

among the sites of which the panelist is a user. In order to capture substitution into dating websites by users who did not previously use these websites, I specify a market for dating websites that extends beyond dating website users as defined above. In particular, I consider panelists who visit a dating website at least once but do not qualify as a user under the criteria above as having chosen my model's outside option for the purposes of estimation; I explain this outside option in detail later in the paper. I drop all other panelists from the sample used in my estimation procedure and counterfactual analysis. After determining whether each panelist is a dating website user, I drop all markets from my analysis with under 100 observed users.

2.1 Multihoming

Consumers in my setting are able to use multiple websites, although few consumers multihome in practice. Among all panelists who use at least one site according to the criterion established by the preceding section, 81% use only one website, 17% use two, and 2% use over two sites. Table 2 reports shares of each site's users who use other dating websites. Not only do few panelists multihome, panelists who multihome generally spend a large share of their time on a single dating website: the average number of minutes that a multihoming panelist spends on the panelist's primary site is 1023, whereas the average number of minutes spent on other sites is only 158. In addition, the average share of time that a panelist spends on the panelist's primary site among all dating websites that the panelist uses is 79%; the median across panelists is 81%. Although multihoming can play an important role in platform competition, the fact that panelists in my sample generally concentrate their online dating activity on a single site motivates my decision to use a model in which each consumer selects a single primary site.

2.2 Relationship between local population and dating website usage

Network externalities could make dating websites 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 reflect that more populous areas could have more offline dating opportunities. To obtain suggestive evidence of these hypotheses, I regress measures of dating

Table 3: Distribution of local population measure

τ	τ^{th} quantile
0.10	8150
0.25	26992
0.50	96418
0.75	226322
0.90	429453

website usage on a measure of local population. I define a consumer's local population as the combined population of ZIP Code Tabulation Area whose geographic centres are within five miles of the geographic centre of the consumer's own ZIP Code Tabulation Area. Table 3 reports some of the quantiles of this variable in my sample of Comscore panelists in 2007–2008. Table 4 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 4 are not estimates of the effect of population on dating website usage given that residents of high population areas are likely unobservably different than people in less populous areas in ways that could affect dating website usage. With that 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 that compete with online dating.

Last, Figure 2 provides evidence of cross-market variation in site market shares. In the presence of network externalities, a site's popularity in a region could be self-fulfilling in that a high number of users on a site explains the site's appeal to these users. In a market with the same exogenous characteristics, a site's lack of popularity could similarly be self-fulfilling. Thus, network externalities could lead a site to be popular in some regions but not others even in the absence of differences in site characteristics.

Table 4: Local population and dating website usage

Panel A: Baseline results

	Usage indicator	Duration	Pages viewed	Sessions
Pop.: under 0.10q	0.0574	9.18	11.24	1.21
	(0.0020)	(0.47)	(0.68)	(0.04)
Pop.: 0.10q to 0.25q	0.0633	9.69	12.50	1.25
	(0.0017)	(0.38)	(0.55)	(0.03)
Pop.: 0.25q to 0.50q	0.0702	11.02	15.14	1.43
	(0.0013)	(0.29)	(0.43)	(0.03)
Pop.: 0.50q to 0.75q	0.0682	10.82	15.43	1.36
	(0.0013)	(0.29)	(0.43)	(0.03)
Pop.: 0.75q to 0.90q	0.0678	10.08	14.24	1.34
	(0.0017)	(0.38)	(0.55)	(0.03)
Pop.: over 0.90q	0.0607	9.35	12.75	1.26
	(0.0020)	(0.47)	(0.68)	(0.04)
R^2	0.066	0.033	0.029	0.066
N	147092	145245	145245	145245

Panel B: Results with controls

	Usage indicator	Duration	Pages viewed	Sessions
Pop.: under 0.10q	0.0679	12.64	12.60	1.29
	(0.0042)	(0.97)	(1.40)	(0.09)
Pop.: 0.10q to 0.25q	0.0728	13.23	13.62	1.30
	(0.0041)	(0.94)	(1.37)	(0.08)
Pop.: 0.25q to 0.50q	0.0787	14.61	16.02	1.46
	(0.0041)	(0.93)	(1.35)	(0.08)
Pop.: 0.50q to 0.75q	0.0769	14.54	16.38	1.39
	(0.0041)	(0.93)	(1.36)	(0.08)
Pop.: 0.75q to 0.90q	0.0765	13.80	15.28	1.36
	(0.0042)	(0.96)	(1.40)	(0.09)
Pop.: over 0.90q	0.0691	13.09	14.00	1.27
	(0.0043)	(0.98)	(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

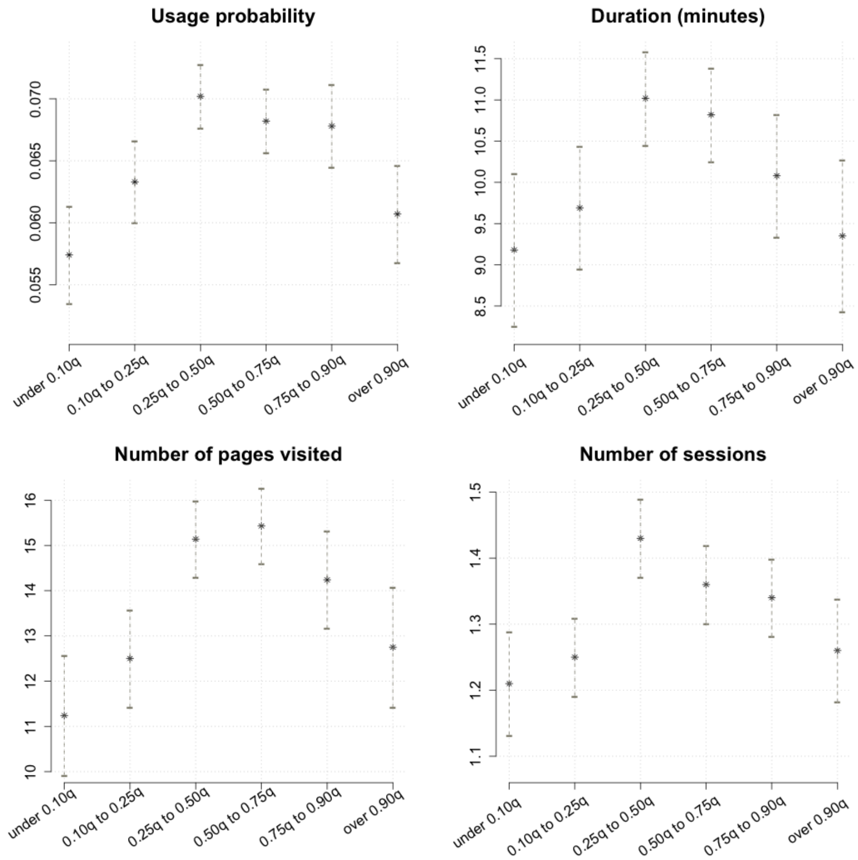
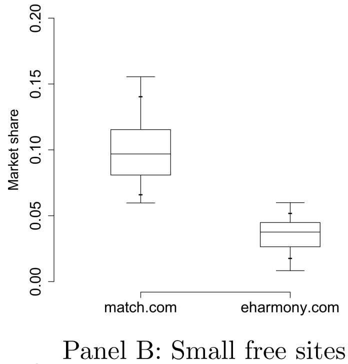
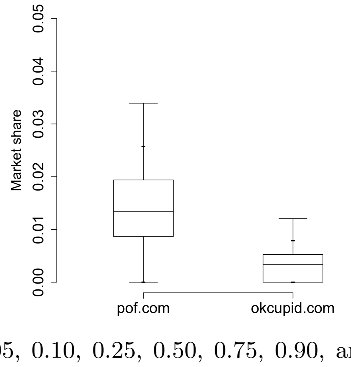


Figure 2: Cross-market variation in sites' market shares in 2007–2008

Panel A: Large paid sites



Panel B: Small free sites



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 Model

This section proposes a semi-nonparametric model of demand with network externalities that is largely based on the general demand model of Berry and Haile (2022). The section also describes a model specification that is especially relevant for empirical applications. Although I often frame the model as a discrete-choice model of demand, this framing is not necessary: consumers choice probabilities can be interpreted as quantities purchased as opposed probabilities of choosing one unit of a specific alternative.

Table 5: Summary of notation

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

The model features markets t, each with a measure M_t continuum of consumers. Each consumer chooses between J products and an outside option. Let $\mathcal{J}$ denote the set of products *excluding* the outside option. Each consumer i belongs to one demographic group d_i among D such groups, and the measure of consumers in demographic group d in market t is M_t^d. In the dating website example, the demographic groups could be defined according to age (i.e., under and over 35 years old) to allow preferences for other users to depend on the age of those users. Consumers i choose between alternatives based on observable individual characteristics w_{it}, observable market characteristics x_t, products' unobservable qualities $\xi_t = \{\xi_{jt}^d : 1 \le j \le J,\ 1 \le d \le D\}$ of products in market t that may vary across demographic groups,[1] and market shares $s_t = \{s_{jt}^d : 1 \le j \le J,\ 1 \le d \le D\}$. 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 $d \in \{1, \ldots, D\}$. Note that, in the market data setting, the researcher observes only market shares s_t and market characteristics x_t. I assume that there is no observed demographic variation in the market data setting and therefore set $D = 1$ when analyzing this case in the main text. In the microdata setting, we additionally observe w_{it}, d_i, and $\mathscr{s}_{ijt}$, the last of which is consumer i's probability of choosing product j conditional on w_{it}, d_i, x_t, ξ_t, and s_t. See Table 5 for a summary of the notation and for a specification of the variables present in market level data and in microdata.

Consumer choice models typically feature prices as product characteristics contained in x_t. Additionally, price is usually thought to be endogenous in the sense that it is dependent on

[1] Although I refer to the ξ_t as qualities, they may reflect unobserved demand shifters that are not best interpreted as product quality, e.g., informative advertising.

ξ_t. In what follows, I ignore price endogeneity (and, more generally, endogenous product characteristics other than s_t) to focus attention on network externalities.

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

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

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

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

The conditioning on t and the t subscript on $\bar{\sigma}_{jt}$ are included to indicate the dependence of average choice probabilities on the distribution of (w_{it}, d_i) on t (i.e., that markets have different distributions of consumer characteristics). In what follows, I use F_t^w to denote the distribution of $\{w_{ijt}\}_{j \in \mathcal{J}_t}$ in t.

Before discussing the determination of market shares s_t, I describe a semi-parametric specification of the general model described above. In this specification, consumer i selects the alternative j that maximizes the indirect utility

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

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

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

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

Although the model features a choice of one alternative, it can accommodate multihoming via the definition of alternatives representing the choice of multiple platforms. For illustration, suppose that a consumer chooses whether to join each of two platforms, and that the consumer receives utility proportional to the share of consumers available through these platforms. Assume that there is only one demographic group for simplicity. In this case, we can let $j = 3$ denote the option of joining both platforms and set $f_1(s_t) = \gamma \times (s_{1t} + s_{3t})$, $f_2(s_t) = \gamma \times (s_{2t} + s_{3t})$, and $f_3(s_t) = \gamma \times (s_{1t} + s_{2t} + s_{3t})$, where γ is the proportionality factor that governs the strength of network externalities.

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

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

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

3.1 Multiple equilibria

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

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

as long as the following matrix is nonsingular:

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

where $D_s\bar{\sigma}_t(\chi_t, s_t)$ is the derivative of $\bar{\sigma}_t(\chi_t, s_t)$ with respect to s_t. I call the unique s^* function defined in the neighbourhood of a particular χ_t an *equilibrium surface* around (χ_t, s_t). Local uniqueness of the equilibrium s_t at market characteristics χ_t means that there is a unique equilibrium surface defined around χ_t. In Appendix B, I discuss the condition that (4) must be nonsingular, and I argue that nonsingularity occurs only in knife-edge cases.

As argued by Bayer et al. (2004), the concept of local uniqueness is useful because it allows for coherent statements about the effects of marginal changes in market characteristics. Indeed, when an equilibrium surface exists around an equilibrium, the slope of this surface with respect to various market characteristics provides marginal effects of changes in market characteristics.

3.2 Microfoundation for network externalities in the dating industry

The presence of network externalities in demand for dating websites can be microfounded using a model of search and matching within websites. Appendix A proposes such a model that gives rise to a dependency between the number of users on a platform and a user's valuation of the platform. This model is a variant of Smith (2006)'s model of search and matching in the marriage market featuring quadratic search technology, exogenous matches and match destruction occurring at exponential rates, and users who idiosyncratically value matches with other users. Appendix A describes the model in detail and shows that a consumer's value from using the platform is increasing in the share of other consumers using the platform.

4 Identification

Although models of demand for differentiated products are identified with market data alone under an appropriate index assumption, the availability of suitable instrumental variables, and other reasonable assumptions—see Berry and Haile (2014)—adding network externalities to these models generally introduces a requirement for microdata and additional assumptions. This section 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 a semi-nonparametric model.

Before beginning my identification analysis in earnest, I clarify the goal of this analysis. I seek to identify the average choice probability function $\bar{\sigma}_t$ in the market data setting and the choice probability function σ in the microdata setting. These functions provide consumers' demanded quantities under various market characteristics and market shares. An alternative object of identification is the mapping from market characteristics $\{x_t, \xi_t\}$ to market shares s_t that is implicitly defined by the fixed point condition (3). This mapping may be identified even when $\bar{\sigma}$ is not identified. In this case, the overall effects of x_t on demand are uniquely determined, but the extent to which these effects owe to preferences for x_t as opposed to network externalities is not uniquely determined. My identification analysis seeks to identify $\bar{\sigma}$ rather than just the mapping $\{x_t, \xi_t\} \mapsto s_t$ for several reasons. First, the welfare consequences of changes in market characteristics depends on the contribution of network externalities to the effects of market characteristics. Consider, for example, decreasing a platform's membership price for a subset of consumers. Absent network externalities, ineligible consumers are no better or worse off due to the targeted price decrease. Under positive network externalities, however, inframarginal ineligible consumers are better off because the targeted price decrease expands the platform's usership. Second, the effects of counterfactual changes in the market environment often depend on the nature of network externalities. The integration of two platforms that combines their user bases, for example, has a greater effect on consumers' willingnesses to pay for the integrated platform when network externalities are stronger. Last, the identification of network externalities enables an evaluation of whether a market is prone to tipping toward a single dominant platform on account of these externalities.

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 specify the indirect utility u_{it} as

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

Here, x_t is an exogenous characteristic of the dating website, 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 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 see 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

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

or, inverting the strictly increasing distribution function F,

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

Suppose that we observe only s_t and x_t for a population of markets t; that is, suppose 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

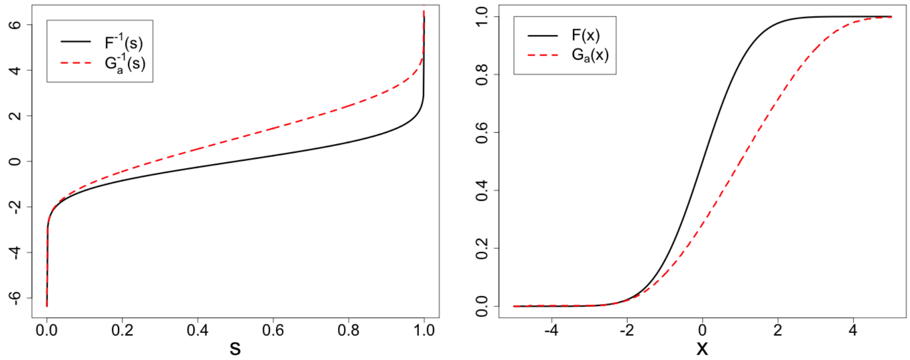
$$G_\delta^{-1}(s) = F^{-1}(s) + \delta s.$$

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 relatively large when δ is low), and

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

Figure 3: Illustration of G_δ



The source of the identification problem is that we cannot fix s_t while varying x_t to identify the effect of x_t on choice probabilities. Note that a researcher who ruled out network externalities when $\gamma > 0$ would overstate the effect of x_t on consumer 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 consumer choices and characteristics allows us to identify the model. To see why, suppose that the individual characteristic w_{it} that we observe is a component of the idiosyncratic taste term ε_{it}: $\varepsilon_{it} = w_{it} + \tilde{\varepsilon}_{it}$. In the dating website setting, the consumer characteristic w_{it} could be an inverse measure of internet speed. Suppose additionally that the remaining unobservable aspect of idiosyncratic tastes $\tilde{\varepsilon}_{it}$ is distributed according to $\tilde{F}$ independently of w_{it} so that

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

Here, δ_t is the average taste for the dating website in market t. We now see that, when y_{it} is an indicator for whether consumer i uses the dating website and the t subscript on $\Pr_t$ indicates conditioning on all characteristics of market t,

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

The left-hand side is observable and the final expression on the right-hand side is one minus the distribution function of $\delta_t + \varepsilon_{it}$. When w_{it} has a large support, (6) 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 $\tilde{F}$.

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 mechanically dependent because ξ_t partly determines s_t. If w_{it}'s distribution F_t^w varies across markets, then functions of this distribution (e.g., its mean) will shift 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.[2] It follows that, as long as x_t is mean-independent of ξ_t (i.e., $\mathbb{E}[\xi_t|x_t] = 0$), each of β, γ, and ξ_t will be identified by an instrumental variables argument.

The argument of Berry and Haile (2022) can be applied to identify this simple model even without a large support assumption. Their approach involves the assumption of a common choice probability s^* that can be achieved in any market t with the appropriate choice of w_{it}. In my setting, the common choice probability condition requires that there is a s^* such that for all t, the support of w_{it} includes a point $w_t(s^*)$ satisfying

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

As long as $\tilde{F}$ is injective, the left-hand side of the equation above can be inverted to obtain

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

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

I now provide an example to build intuition for the identification strategy 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 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 (6). 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 instrumental variables approach described above could fail. First, the approach is threatened if consumers locate based on ξ_t. As an example, consumers

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

who enjoy using the dating website may locate in areas where fast internet is more readily accessible, which would induce dependence between average internet speeds and unobserved tastes for dating website. 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}, then ξ_t will depend on F_t^w. Contextual network effects also threaten the instrumental variables approach. The literatures on peer effects and network externalities use the term *contextual network effects* to describe the case in which an agent's outcome of interest (e.g., choice of platform) depends directly on the characteristics of connected agents (e.g., people in the same market). For the market-specific distribution F_t^w of w_{it} to directly affect u_{it} in my model, it must enter through the unobservable ξ_t. This induces a violation of the identification condition. The identification threat posed by contextual network effects means that the researcher must be careful in selecting which characteristics w_{it} to use in constructing instruments.

In the binary choice setting above, instruments based on F_t^w are the only excluded shifters of s_t available. In a setting with multiple alternatives, characteristics of other alternatives—which are called BLP instruments in the literature on differentiated products demand markets following their use by Berry et al. (1995)—become available under standard exclusion restrictions. These additional instruments, however, are not available in settings with multiple alternatives without cross-market variation in the choice set. This is relevant because platforms competing across different geographically defined markets often exhibit little or no variation in observable characteristics across regions. Therefore, the previous paragraph's enumeration of plausible ways for instruments based on F_t^w to fail may seem pessimistic about prospects for identifying demand with network externalities. Solutions include the randomization of characteristics w_{it} (e.g., through platform-adoption subsidies) or the collection of data on individual characteristics that are unlikely to drive targeted promotion or contextual network effects.

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, M_t and s_t will be uncorrelated. When there are network externalities, 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, $\gamma < 0$). This approach, however, requires the strong assumption that M_t is excluded from u_{it} apart from the $M_t s_t$ term; this assumption is violated when tastes for platforms differ across markets of different sizes conditional on market characteristics x_t.

4.2 Identification with market data

The remainder of this section formalizes the identification arguments presented in the preceding subsection; a reader who is uninterested in the technical analysis of the identification of demand with network externalities may proceed to Section 5. I begin by considering the identification of a demand model with market data under an index structure similar to that of Berry and Haile (2014):

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

where $x_t = [x_{1t}, \ldots, x_{Jt}]'$ is a vector of product characteristics that vary across markets. I focus on the average choice probability function $\bar{\sigma}$ rather than the consumer-level choice probability function σ as the object that I seek to identify with market data because the consumer characteristics that enter σ_{ijt} as arguments are not available in market data. Although $\bar{\sigma}$ may depend on t through the distribution of individual characteristics w_{it}, I rule out this dependency when analyzing identification with market data because w_{it} are not observable in this setting. Note also that my approach allows for product characteristics other than x_t, which are required to satisfy the index restriction stated in (8); as Berry and Haile (2014) suggest, the researcher can condition on exogenous characteristics $x_t^{(2)}$ and suppress them in the notation, implicitly identifying demand conditional on each $x_t^{(2)}$ in these characteristics' support under which the assumptions invoked to identify demand hold.

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

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

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

Berry et al. (2013) provide sufficient conditions for Assumption INVERT-MARKET that require a minimal amount of substitutability between products. Berry and Haile (2014) provide identification results under these conditions, which they use to ensure an invertibility property similar to that which I directly assume in Assumption INVERT-MARKET. Assumption

INVERT-MARKET implies the uniqueness of a vector $\bar{\sigma}^{-1}(\delta)$ that solves $\bar{\sigma}^{-1}(\bar{\sigma}(\delta)) = \delta$ for δ in the support of $\delta(x_t, s_t, \xi_t)$.

The now provide the primary negative result of this section.

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

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

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

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

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

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

The simple proof of Proposition 1 demonstrates the nature of the identification problem: market shares appear in the inverse demand equation (9) directly through $\bar{\sigma}^{-1}$ and also through the network externality function h. This problem reflects that substitution patterns and network externalities cannot be empirically disentangled. To illustrate, an improvement in a product that increases its market share at the expense of an alternative product could be explained by either (i) a high degree of substitutability between the products, or (ii) both products being subject to strong network externalities that amplify the first product's gains from its improvement and the second product's losses. Note that the identification problem does not relate to the unavailability of instruments: even if s_t were mean independent of ξ_t, which is generally impossible, the model's primitives would be unidentified.

Under the availability of instruments satisfying an appropriate exclusion restriction and completeness condition, ξ_t is identified from the nonparametric regression equation $x_{jt} = \bar{\sigma}^{-1}(s_t) - h_j(s_t) - \xi_{jt}$, where $\bar{\sigma}^{-1}(s_t) - h_j(s_t)$ is the corresponding nonparametric regression function. Although the identification of ξ_t implies the identification of the mapping $\{x_t, \xi_t\} \mapsto s_t$, it does not imply the identification of $\bar{\sigma}$, which is the object of interest in my identification analysis.

Identification problems exist even when $\bar{\sigma}$ is known. Note that the assumption of a known $\bar{\sigma}$ is invoked when the consumer/choice-level unobservables ε_{ij} in a discrete-choice model are assumed to follow a specific parametric distribution. Such an assumption amounts to imposing substitution patterns between products. In the case of a known $\bar{\sigma}$, the inverse demand equation (9) becomes

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

I introduce the coefficients β_j because setting the coefficient of each x_{jt} to one is no longer a scale normalization when $\bar{\sigma}$ is known. Note that there are J included endogenous regressors s_t and only $J-1$ distinct instruments that are plausibly exogenous and capable of shifting s_t in the context of the model, namely the BLP instruments. The deficit of instruments relative to endogenous regressors could be rectified by excluding certain market shares from $h_j(s_t)$ or expanding the dimension of x_{jt} to obtain additional BLP instruments. I do not, however, earnestly recommend making such assumptions; I outline the approach of assuming a known $\bar{\sigma}$ primarily to detail the woes of identifying a demand model with network externalities with market data.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Proof. See Appendix C.2. □

Proposition 2 establishes that the prospects for identification improve when choice probabilities depend on total quantities rather than market shares, although the proposition's conditions for identification are fairly strong. The "large support" assumption is useful because it facilitates an identification-at-infinity argument relying on network externalities becoming irrelevant as the market size tends to zero. Note that neither assumptions (a) nor (b) is required to identify ξ_t, which is sufficient for the identification of the mapping from $\{x_t, \xi_t\}$ to s_t. Assumptions (a) and (b) instead provide two different ways of separating the effects of market characteristics on market shares into direct effects and indirect effects owing to network externalities.

Appendix C.1 provides analysis of identification of the total quantities model in the case in which products' own quantities enter as additively separable linear terms in their δ indices. This functional restriction permits identification of the model.

4.3 Identification with microdata

The identification of the model with microdata follows from the argument of Berry and Haile (2022) with only minor adjustments. The fact that the model is identified under reasonable assumptions with microdata but not with market data reflects the fact that, as Berry and Haile (2022) note, microdata eliminates the need for instruments for quantities in a setting without network externalities. Berry and Haile (2014) and Berry and Haile (2022) identify inverse demand using nonparametric instrumental variables equations in the market data setting and the microdata setting, respectively. In the microdata case, market shares do not directly enter this inverse demand function, which eliminates the need for instruments for these variables. The elimination of a need for market share instruments reflects that within-market variation allows for the identification of substitution patterns. Introducing network externalities reintroduces market shares as arguments of this function, but instruments may be available to nonetheless identify demand. This is impossible in the market data case because market shares enter the inverse demand function in two separate ways: directly in a sense that reflects inverse demand is the inverse of a mapping into quantities (this is the mapping that governs substitution patterns) and indirectly through the dependence of demand on network externalities.

I impose the following index structure on consumer i's choice probability function:

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

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

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

for $j \in \{1, \dots, J\}$. I additionally impose that the individual characteristics w_{it} have dimension of at least J. The identification argument that follows does not make use of market characteristics x_t; therefore, I condition on these characteristics and suppress them in the notation. Additionally, the argument requires individual characteristics w_{it} of dimension J. Other individual characteristics may enter the model, but they do not play an important role in the identification analysis and are therefore omitted in my exposition.[3]

I now provide assumptions are adapted with little revision from Assumptions 2 and 3 in Berry and Haile (2022).

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

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

The model requires several normalizations to achieve identification. First, I impose $\mathbb{E}[\xi_t] = 0$ to rule out nonidentification due to the shifting of a constant between ξ_t and g. Second, I impose that $g(w_0) = 0$ for a known value $w_0 \in \operatorname{supp} w_{it}$. As a scale normalization, I impose $Dg(w_0) = I$, where D denotes the derivative operator. These last two normalizations are analogous to equations (7) and (8) in Berry and Haile (2022).

I additionally require the availability of variables z_t that satisfy the validity and relevance

[3]See Berry and Haile (2022) for a treatment of additional individual characteristics, which they denote by Y_{it}.

conditions required of instruments for market shares. These assumptions are the standard exclusion restriction and completeness condition of Newey and Powell (2003).

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

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

Work on the identification of discrete-choice models often assumes that vectors of so-called "special regressors" are supported on the entirety of the relevant Euclidean space. Berry and Haile (2022) use an alternative and less restrictive assumption that serves a similar purpose in their identification proof. Assumption CPROB adapts this assumption to my setting. Additionally, Berry and Haile (2022) impose a nondegeneracy condition on the demand unobservables ξ_t that I adapt in Assumption NOND.

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

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

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

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

The identification argument underlying Proposition 3 is nearly identical to that in Berry and Haile (2022) with the exception that it involves instrumenting for s_t rather than an endogenous product characteristic. My preferred way to construct the instruments z_{jt} appearing in Assumptions NPIV-EX and NPIV-C is to use cross-market variation in distributions of consumer characteristics w_{ijt}. One example is $z_t = \mathbb{E}_t[w_{it}]$, the mean value of w_{it} in market t. In the context of the semi-parametric specification outlined by Section 3, another candidate instrument vector $z_t = [z_{1t}, \ldots, z_{Jt}]$ has components $z_{jt} = \Pr_t(w_{it}'\lambda_j + \varepsilon_{ijt} = \max_k w_{it}'\lambda_k + \varepsilon_{ikt})$ equal to the predicted shares of products j in market t when all market-level shifters (i.e., x_{jt}, s_t, and ξ_t) are removed from the utility equation. As long as F_t^w is 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 instruments. See Section 4.4 for additional discussion of available instruments.

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

4.4 Discussion of identification results

I now discuss my identification results in comparison to related research. Bayer and Timmins (2007) consider the estimation of a discrete-choice model with network externalities by applying an instrumental variables approach to microdata. In their model, consumer choice probabilities depend on market shares rather than total quantities. Timmins and Murdock (2007), who follow the approach described by Bayer and Timmins (2007), use BLP instruments in estimating network externalities. This approach is suitable for the empirical setting of Timmins and Murdock (2007) 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 the approaches used by Bayer and Timmins (2007) and Timmins and Murdock (2007) under 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 instrumental variables used by Guiteras et al. (2019), who model demand for latrines in Bangladeshi villages using a discrete-choice model with network externalities, They estimate this model with data from a randomized control trial that provided latrine subsidies to a random subset of households in their sample that included randomization across neighbourhoods and villages in subsidy eligibility rates. Guiteras et al. (2019) use the share of households eligible for subsidies in a neighbourhood as an instrument for the share of households buying latrines in that neighbourhood. These instruments are examples of the instruments based on market-specific distributions of consumer characteristics that I consider in my identification analysis. My paper expands upon the identification insight of Guiteras et al. (2019) by showing that demographic characteristics not directly related to a characteristic of the analyzed product can aid in identifying network externalities.

Instruments based on market-specific distributions of consumer characteristics, which I emphasize in my identification analysis and which are used by Guiteras et al. (2019), are coined *Waldfogel instruments* by Berry and Haile (2016) in response to the insights of Waldfogel (2003) and Waldfogel (2008) that the local demographic profile influence consumers' choice sets. As noted in Section 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, (ii) consumers choose their geographical markets based on the local quality of their preferred products, and (iii) the presence of contextual network externalities by which the local demographic profile directly affects tastes for products.[4]

[4]Another challenge associated with Waldfogel instruments is that they become weaker as the researcher

My analysis of models in which total quantities rather than market shares enter choice probability functions shows that these types of models have similar identification properties. One difference is that, in the total quantities model, the market size is available as an instrumental variable under a standard exclusion restriction. This exclusion restriction is strong: it requires that consumer tastes do not differ between markets of different sizes (e.g., large cities versus small towns) conditional on other observable market and consumer characteristics. Another difference is that cross-market variation in market size can separate the direct and indirect effects of market characteristics under plausible buy strong assumptions. This is true even in the market data setting. Proposition 2 provides these strong assumptions, one of which is the case in which we can take a limit as the market size goes to zero. In this case, network externalities and the indirect effect mentioned above vanish, which allows the direct and indirect effects to be separately identified.

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. Although estimating all parameters simultaneously may introduce efficiency gains, such an estimator is likely to be difficult to compute in practice. The estimator that I outline, in contrast, involves two simple estimators (a maximum likelihood estimator of a logit model and a two-stage least squares estimator) that do not entail a serious computational burden.

5.1 Asymptotic assumptions

Before outlining my estimation procedure, I provide notation and asymptotic assumptions under which I conduct inference. I write the data as

$$D^k = \{D_{it} : 1 \leq i \leq N_t^k, 1 \leq t \leq T^k\},$$

where k is an index that tends to infinity in my asymptotic analysis and D_{it} is a vector containing all observables pertaining to individual i in market t. I assume that, within each market t and for all k, $D_{it} \sim_{\text{iid}} \Psi_t$ for a distribution Ψ_t. My primary asymptotic assumptions

specifies increasingly narrow demographic groups. Waldfogel instruments rely on variation within subpopulations of markets belonging to particular demographic groups to shift market shares within these subpopulations. Defining narrow demographic groups limits variation within these subpopulations. To fix ideas, consider a case in which there are two demographic groups, call them young and old. 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. Therefore, the Waldfogel instrument computed as the proportion of young consumers in the market is irrelevant in this case.

are

$$T^k \xrightarrow{k\uparrow\infty} \infty$$
$$N_t^k / T^k \xrightarrow{k\uparrow\infty} \infty.$$

That is, I assume that the number of markets T^k tends to infinity and the number of observations within each market, N_t^k, also tends to infinity. This latter assumption is reasonable in settings such as my own in which the number of individuals in the microdata within each market is large. Last, I assume that the number of platforms J is fixed (i.e., it does not depend on k).

5.2 Microstep

The estimating equation of the microstep is

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

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

5.3 Market step

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

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

See Section 3 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 in estimates of the market shares computed using the Comscore data for the true market shares. The x_{jt} are firm-time period indicators whose coefficients I denote by ψ_j so that $x_{jt}'\beta = \psi_j$. Last, I estimate (14) using several different parametric forms of the network externality function $f_j^d(\cdot;\gamma)$, as I state explicitly in Section 6.

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 (14), I compute the predicted value of the market share based on the microstep estimates when (i) $\delta_{jt}^d = 0$ for all j, t, d and (ii) certain effects of individual characteristics w_{ijt} are also set to zero. These predicted values, which I use as instruments for the estimation of network externalities, are examples of the Waldfogel instruments discussed by Section 4.4. I apply the same transformation to the predicted market shares when constructing the instruments as I do to the market shares when entering them into (14). To illustrate, 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

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

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

In practice, I use estimates of the platform/market utility indices δ_{jt} from the microstep of my estimation procedure rather than the true values of these indices in estimating (14). Sampling error in estimation of δ_{jt} is asymptotically irrelevant, however, under the maintained assumption that and the number of observations within each market t tends to infinity. I also enter estimates $\hat{s}_{jt}$ in place of the market shares s_{jt} in (14). These estimates, which are empirical choice frequencies from my estimation sample, are also mismeasured on account of sampling error. This measurement error is both asymptotically irrelevant under my maintained asymptotic assumptions, and also rectified in a finite sample by instrumenting s_{jt} with z_{jt}.

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

5.4 Price sensitivity

To this point, I have ignored price competition between firms. Estimating consumer price sensitivity is important for computing pricing equilibria in counterfactuals and for expressing welfare figures in dollar terms. But the fact that the dating websites in my sample charge uniform prices across geography means that I observe minimal price variation, which prevents me from estimating price sensitivity the 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_j in (14) and let ψ_j denote the fixed effect for site j. I then make the decomposition $\psi_j = \bar{\psi}_j - \alpha p_j$ and assume that the observed prices $\{p_j^*\}$ constitute a Bertrand-Nash pricing equilibrium with marginal costs of zero in that

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

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

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

which provided the basis of my estimation of α.

To compute my estimator $\hat{\alpha}$ of α, I substitute empirical analogues/estimates for population objects/parameters in each paid site's FOC (16) and then solve for α. These FOCs include price derivatives of market shares, which are not well defined without an assumption on how prices affect equilibrium selection in the presence of multiple equilibria. I assume that firms believe their market shares at counterfactual prices are given by the equilibrium surface around (χ_t, s_t) as defined in Section 3.1, where χ_t includes firms' prices in market t. I then use the price derivatives of this equilibrium surface as the price derivatives appearing in the FOCs underlying my estimation of α. The implicit function theorem provides an explicit form for these derivatives. 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 reflecting that the direct effect changes product j's market share which in turn affects the network externality term in product j's indirect utility. Some of the sites in my sample are free to use. I do not include these sites' FOCs in the estimation of α and I assume that free sites remain free in my counterfactuals. Each paid site's FOC provides a separate estimate of α; my final estimator $\hat{\alpha}$ is the average of these site-specific estimates. I compute standard errors for $\hat{\alpha}$ using a parametric bootstrap that involves sampling from the estimated asymptotic distribution of the the parameters estimated in the market step of estimation. I include additional details on the estimation of α in Appendix D.

I check whether $\hat{\alpha}$ is a reasonable estimator by considering its implications for a price response to monopolization. Section 7.4 considers a counterfactual in which `match.com` becomes a monopolist. Under estimates of my preferred specifications, `match.com` raises its price by 30.5% upon becoming a monopolist (see Table 16); this magnitude seems sensible.

6 Parameter estimates

This section reports and discusses my parameter estimates. I estimate the model using two different specifications of demographic groups:

(**Overall**) All consumers belong to the same 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 Section 5.3. The sites that consumers choose between are `eharmony.com`, `match.com`, `okcupid.com` and `pof.com`; choices to use other sites and a failure to use any site are grouped together in the outside option. The microstep of estimation involves a large number of parameters whose presentation I relegate to Appendix E; see Tables 18 through 21. Many of the estimated parameters indicate significant taste differences across individuals with different observable characteristics.

Panel A of Table 6 displays the parameter estimates of an share-type model with the "Overall" demographic specification and the network externality function specification $f_j(s_t; \gamma) = \gamma \log(s_{jt})$. Instrumenting for market shares with the demographic instruments decreases the estimated coefficient of the network externality term relative to OLS. This reflects the 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") provide the estimated site intercepts.

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

Tables 7 reports estimates from the "Age" demographic group specification under various specifications of the network externality function. In particular, column (1) of each table reports estimates from a specification in which consumers care only about the market share of a site within their own demographic group; column (2) reports estimates from a specification in which a consumer's tastes depend on the market share of a site within their own demographic group and the other demographic group; column (3) reports estimates from a specification in which a consumer's tastes depend on the market share of a site within their own demographic group only, but members of different demographic groups have different preferences for their own-group market shares; and column (4) reports estimates from a specification in which a consumer's tastes depend on the market share of a site both within their own demographic group and within the entire population. These tables suggest considerable homophily within groupings defined by age, as consumers more highly value market shares within their own age group than shares within the other group. Column (3) suggests that the

Table 6: Market step parameter estimates – "Overall" demographic group specification

Panel A: Parameter estimates

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

Panel B: First stage of IV regression

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

strength of network externalities is similar within each demographic group—i.e., that there is not considerable heterogeneity in tastes for market shares between groups—although the own-group tastes for the older age group are not precisely estimated.

We can use the results presented above to compute the value of an increase in a site's usership to an inframarginal of the site in dollar terms. The estimates for the market shares model presented in Table 6 imply that a 10% increase in the usership of a site is worth \$6.34 a month to an inframarginal user of that site whereas the estimates for the total quantities model presented in Table 8 imply that a 10% increase in the usership of a site is worth \$6.67 a month to an inframarginal user of that site. Given that `match.com`'s monthly subscription price in 2007 was \$34.99, a 10% increasing in `match.com`'s usership in a particular market would be worth 18% of the site's price to an inframarginal `match.com` user under estimates of the market share model and 19% under estimates of the total quantities model.

7 Counterfactual analysis

I use my estimated model to decompose cross-market variation in market shares and to assess the effects of changes in the market structure of the dating websites industry. The model estimates that I use in this analysis are those for the model without distinct demographic groups that include market shares rather than total quantities as utility shifters; see Tables

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

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

Table 8: Market step parameter estimates – "Overall" demographic group specification, quantity-type model

Panel A: Parameter estimates

	OLS	IV
$\log(M_t s_{jt})$	0.98	0.69
	(0.02)	(0.14)
$\log(M_t)$	-0.98	-0.72
	(0.02)	(0.12)
eharmony	0.07	-0.23
	(0.16)	(0.28)
match	-0.84	-0.92
	(0.16)	(0.24)
okcupid	-3.41	-4.31
	(0.18)	(0.49)
pof	0.41	-0.13
	(0.17)	(0.35)
p_j $(\hat{\alpha})$		0.0099
		(0.0043)

Panel B: First stage of IV regression

	$\widetilde{\log(s_{jt})}$
$\tilde{z}_{jt}$	0.97
	(0.32)
F	9.12

18 for the microstep estimates and 6 for the market step estimates. I compare outcomes in equilibria under counterfactual market environments with outcomes in equilibria under the baseline market environment, i.e., the observed non-price website characteristics and market structure. The prices in the baseline equilibrium to which I compare counterfactual equilibrium constitute a Bertrand-Nash pricing equilibrium; I do not impose the observed prices on this baseline equilibrium. Given that each dating website charges a single nation-wide price in my data, I constrain each website's price to be constant across markets when computing equilibria.

I consider how changes in market structure affect consumer welfare but on website profitability throughout this section. As I elaborate upon in a subsequent section, the two primary effects on consumer welfare that I consider are *price response effects* and *network externality effects*. The former are defined as dollarized differences in expected utility between counterfactual equilibria with and without price responses. The latter are defined as differences in consumer enjoyment of network externalities between counterfactual equilibria with price responses and baseline equilibria. Overall differences in expected consumer utility between baseline and counterfactual equilibria also depend on changes in the realized ε_{ijt} idiosyncratic taste shocks of consumers' realized choices. Some of the counterfactuals that I consider remove websites from the market, which mechanically lowers expected utility in a way that is highly dependent on the assumed distribution of the ε_{ijt} shocks. Other counterfactuals adjust consumer utilities for websites; these adjustments also mechanically affect welfare. To avoid these two sorts of mechanical effects from influencing my headline welfare results, and to highlight the roles of price responses and network externalities, I measure overall welfare effects using the sum of the price response effects and network externality effects rather than differences in expected utility.

I compute the equilibria described in this section using a nested iterated best response algorithm. This algorithm is nested in the sense that it (i) iterates on market shares in an inner loop to find a fixed-point of the market share function for a particular vector of prices and (ii) then iterates on prices in an outer loop to find a price vector such that each firm's price maximizes its own profits given the other firm's price.

The first counterfactual analysis that I consider facilitates a decomposition of cross-market variation in 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 Section 3.1, it is generally difficult to compute the effects of counterfactual changes in market characteristics when there 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 Section 3.1 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 Section 3.1) with respect to market characteristics. In practice, 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 characteristics on these outcomes.

I also study the effects of making `match.com` a monopolist. 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 does not arise when there is only one site.

7.1 Decomposition of variance in market shares

I decompose the variation in market shares across markets into three sources: network externalities; unobserved differences in taste across markets; and observed differences in tastes across markets reflecting demographic differences. Table 9 displays the results of this decomposition. For each regime, I compute the standard deviation of market shares across geographies. 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 net. ext." column, I remove the network externality function from consumers' indirect utilities and re-compute market shares in each market. Note that eliminating the network externality function from 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. 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 9 tells us that network externalities explain most of the variation in market shares across geographical markets. Indeed, 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.

Table 9: Decomposition of geographic variation in market shares

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

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. I focus on welfare in the eight most populous markets in my data; Table 10 gives the market shares of the two most popular dating websites, `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.

To adjust the values of `match.com` and `eharmony.com`, I increase each δ_{jt} index for `match.com` by $2 \times \hat{\alpha}$ and decrease each δ_{jt} index for `eharmony.com` by the same amount. Adjusting these indices mechanically changes the utilities enjoyed by sites' inframarginal consumers absent any responses by sites or consumers. To ensure that my welfare results do not reflect these mechanical effects, I present only the welfare changes from price responses and from changes in realized network externalities. I define the *price response effect* as the difference in expected consumer utility between (i) a counterfactual equilibrium in which sites adjust their prices from the baseline equilibrium and (ii) a counterfactual equilibrium with prices fixed at their baseline levels. I define the *network externality effect* as the expected change in the network externality term in the indirect utility of the consumer's selected site between the counterfactual equilibrium with price adjustment and the baseline equilibrium. In the above, expected consumer utility refers to the expectation of $\max_j u_{ij}$ over both ε_{ij} and consumer characteristics w_{ij}. Also, when the consumer selects the outside option, I take it that the network externality term for the consumer's choice is zero. I measure the *net effect* of the counterfactual change on expected utility as the sum of the price response effect and the network externalities effect. Note that this net welfare change does not capture changes in the expected logit shock of the consumer's selected platform.

Table 11 reports the distribution of counterfactual market changes. In particular, it displays various quantiles of the distribution of market-share changes (in percentage points) taken across markets. This table shows that market shares change only marginally in response to the small counterfactual changes in values to consumers, which suggests the counterfactual equilibria belong to the same equilibria surfaces as the baseline equilibria.

Table 12 displays price effects whereas Figure 4 reports welfare effects (in dollars) for several of the largest markets in the analysis. Figure 4 and Table 10 show that, among the most populous markets, the net welfare changes from increasing market concentration are negative.

Table 10: Market shares in top markets under the baseline equilibrium

Market	`match.com`	`eharmony.com`
New York (NY)	0.155	0.135
Los Angeles (CA)	0.146	0.136
New York (NJ)	0.143	0.128
Miami (FL)	0.177	0.132
Chicago (IL)	0.140	0.159
Atlanta (GA)	0.135	0.183
San Jose (CA)	0.165	0.146
Dallas (TX)	0.144	0.138

Table 11: Market share changes in the increase concentration counterfactual

	Quantiles of market share changes (%)				
Site	0.01	0.25	0.50	0.75	0.99
eharmony.com	-0.7433	-0.6067	-0.6067	-0.5360	-0.4495
match.com	0.4210	0.5209	0.5209	0.5923	0.7214
okcupid.com	-0.0092	-0.0038	-0.0038	0.0004	0.0038
pof.com	-0.0142	-0.0038	-0.0038	0.0091	0.0260

These negative effects reflect price increases at `match.com` and, in most markets, decreases in benefits from network externalities suffered by inframarginal users of `eharmony.com`. In Miami, where `match.com` has an especially high baseline market share, the average benefit from network externalities increases in the counterfactual, which partially offsets the negative effect of `match.com`'s price increase on welfare. In Section 7.1, I found that much of the variation in website 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 firm across geography, which in turn implies heterogeneity across geography in the effects of boosting the national market leader on realized network externalities.

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 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.

Table 12: Price changes in the increase concentration counterfactual

Site	Baseline	Counterfactual	Difference
eharmony.com	44.20	43.57	-0.63
match.com	44.66	45.33	0.67

Figure 4: Welfare changes in the increase concentration counterfactual

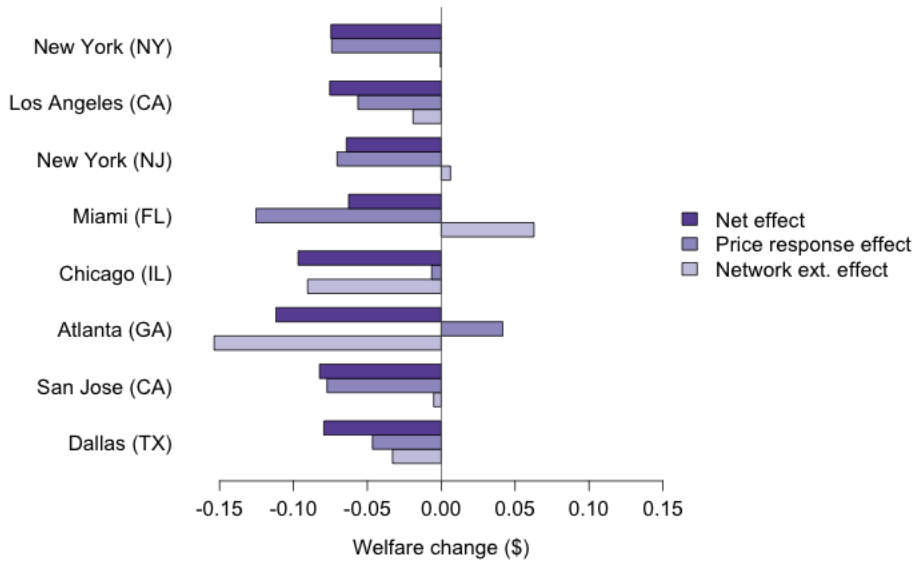


Table 13: Market share changes in the niche site counterfactual

	Quantiles of market share changes				
Site	0.01	0.25	0.50	0.75	0.99
eharmony.com	-0.0235	-0.0149	-0.0149	-0.0087	-0.0023
match.com	-0.0255	-0.0153	-0.0153	-0.0079	-0.0017
okcupid.com	-0.0031	-0.0016	-0.0016	-0.0009	-0.0006
pof.com	0.0209	0.0410	0.0410	0.0648	0.0949

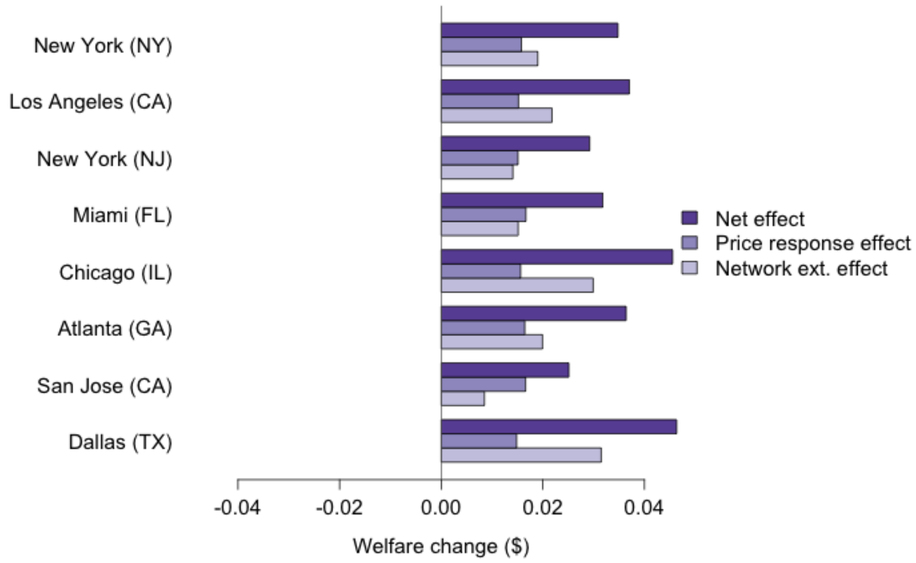
This counterfactual involves marginally increasing (by $10) the appeal of `pof.com` to consumers with college degrees and advanced degrees. In assessing the welfare effects of these changes, I use the measures of price response effects, network externality effects, and net effects that I introduced in the preceding subsection.

Table 13, Table 14, and Figure 5 are the analogues of Table 11, Table 12, and Figure 4 for this "niche site" counterfactual. Increasing the appeal of `pof.com` increases each component of average welfare. Note that utility from network externalities increases 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 suffer 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 educated consumers makes the market for dating websites more competitive and induces `eharmony.com` and `match.com` to lower their prices.

Table 14: Price changes in the niche site counterfactual

	Baseline	Counterfactual	Difference
eharmony.com	44.20	44.18	-0.02
match.com	44.66	44.64	-0.02

Figure 5: Welfare changes in the niche site counterfactual



7.4 Monopoly counterfactual

I now consider a counterfactual in which `match.com` becomes a monopolist. The first two rows of Table 15 compare market shares in this counterfactual to baseline market shares whereas Table 16 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. The reported effects are the same as those introduced in Section 7.2. The results displayed by Figure 6 imply that monopolization decreases average welfare by \$3.00–\$4.53 across the most populous markets. Net welfare losses, which owe to The monopolist website's price increase is responsible for the negative overall effect of monopolization, although consumer welfare losses from this price increases are partially offset by increased enjoyment of network externalities under monopoly.

7.5 Profitability of website integration

I conclude my counterfactual analysis by comparing a scenario with only two dating websites—`eharmony.com` and `match.com`—that are jointly owned but not integrated with a scenario

Table 15: Market share changes in the monopoly counterfactual

	`eharmony.com`	`match.com`	`okcupid.com`	`pof.com`	Outside option
Baseline	0.150	0.147	0.012	0.040	0.651
Monopoly	0.000	0.191	0.000	0.000	0.809
Joint ownership	0.124	0.121	0.000	0.000	0.755

Table 16: Price change for `match.com` in the monopoly counterfactual

Quantity	Value
Baseline	\$44.66
Counterfactual	\$58.28
Change	30.5%

Figure 6: Welfare changes in the monopoly counterfactual

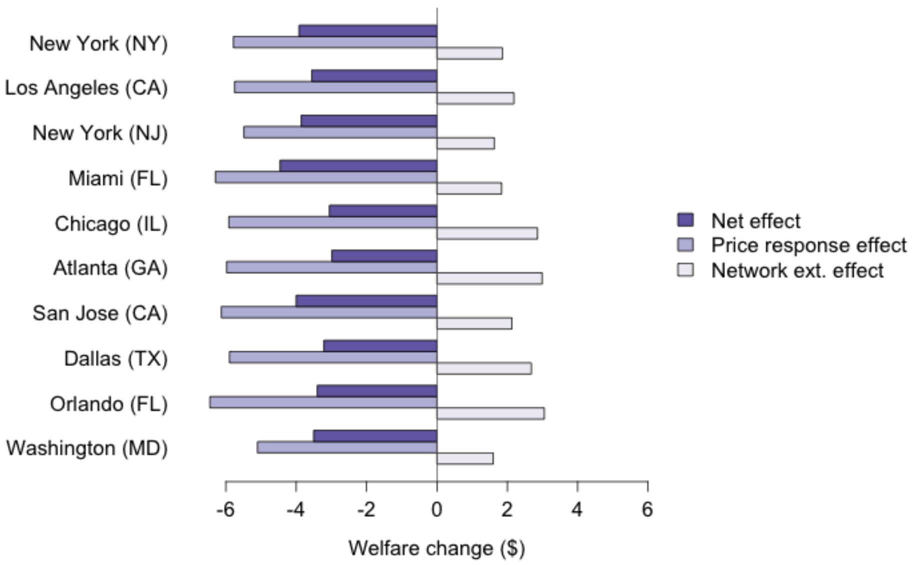


Table 17: Price change in the joint ownership counterfactual

Quantity	Value	Value
Baseline	\$44.20	\$44.66
Counterfactual	\$65.50	\$66.48
Change	48.2%	48.9%

in which `match.com` is a monopolist. The latter scenario corresponds to one in which the joint owner of `eharmony.com` and `match.com` decides to integrate the former website into the latter. Table 15 compares market shares under the joint ownership and monopoly (i.e., integration) scenarios when prices are set to optimize total dating website profits in each case. The sum of market shares for dating websites is higher under joint ownership without integration, although `match.com`'s share is predictably lower under joint ownership. Table 17 compares prices between the two scenarios; it shows that the firm enacts greater price increases under non-integration. This is because a price increase at one website leads consumers to divert to the firm's other website, which mitigates the firm's overall sales losses from price increases. The firm's profits are 31% lower in the monopoly scenario relative to the joint ownership scenario. The fact that the firm's profits fall upon integration partly reflects significant horizontal differentiation of websites, which is evident from the coexistence of several dating websites with sizeable market shares in my data. This differentiation implies that a multi-website firm loses many of a website's users upon abolishing the site: in my counterfactual analysis, the combined market share of the firm's sites falls from 24% to 19%. One benefit to the firm of removing a website is that this removal provides the firm's other website with a higher market share, which in turn increases consumers' willingnesses to pay for this site. Offering differentiated sites, however, allows the firm to support higher prices than a single site with higher market share can command.

8 Conclusion

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

Bibliography

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

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

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

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

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

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

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

Berry, Steven, and Philip Haile. 2022. "Nonparametric Identification of Differentiated Products Demand Using Micro Data." Unpublished working paper.

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

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

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

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

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

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

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

Farronato, Charia, Jessica Fong, and Andrey Fradkin. 2020. "Dog eat dog: measuring network effects using a digital platform merger." NBER working paper 28047.

Fong, Jessica. 2020. "Search, Selectivity, and Market Thickness in Two-Sided Markets: Evidence from Online Dating." Unpublished working paper.

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

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

Guiteras, Raymond, James Levinsohn, and Ahmed Mushfiq Mobarak. 2019. "Demand estimation with strategic complementarities: sanitation in Bangladesh." BREAD Working Paper 553.

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

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

Kim, Jin-Hyuk, Peter Newberry, Liad Wagman, and Ran Wolff. 2021. "Local network effects in the adoption of a digital platform." *Journal of Industrial Economics*, Forthcoming.

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

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

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

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

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

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

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

A Microfoundation for network externalities on dating websites

This appendix provides a model of search and matching on a dating website that justifies the inclusion of a network externality term in consumer payoffs from using dating websites. This model is a variant of that proposed by Smith (2006); the model features a mass L of users of a dating website, each of whom finds matches on the site at an exponential rate ρ. Matches are exogenously destroyed at rate δ. Each user's payoffs are discounted at the interest rate r. The flow value of a match between agents i and j is f_{ij} and the flow value of remaining unmatched is zero. I assume that f_{ij} are identically and independently distributed across pairs (i, j) from the distribution F. Let $\eta = \Pr(f_{ij} > 0)$ and let $\mu_f = \mathbb{E}[f_{ij} | f_{ij} > 0]$. Suppose that consumer i immediately forms any match that yields a positive flow value f_{ij} and does not account for the option value of waiting for a better match. The average present value[5] of being unmatched is

$$V_u = \frac{\rho\eta^2}{r} \int [V_m(f) - V_u] dF(f \mid f > 0), \tag{17}$$

where $u \leq L$ is the mass of unmatched users and $V_m(f)$ is the average present value of being matched with match value f:

$$V_m(f) = f + \frac{\delta}{r}[V_u - V_m(f)]. \tag{18}$$

We can solve (18) to obtain

$$V_m(f) = \frac{rf + \delta V_u}{r + \delta}$$

and then solve (17) to obtain

$$V_u = \frac{\eta^2 u \mu_f}{\psi + \eta^2 u}, \tag{19}$$

where $\psi = (r + \delta)/\rho$ is a measure of the search frictions in the market. The steady-state value of u is given by the condition

$$\delta(L - u) = \rho\eta^2 u^2,$$

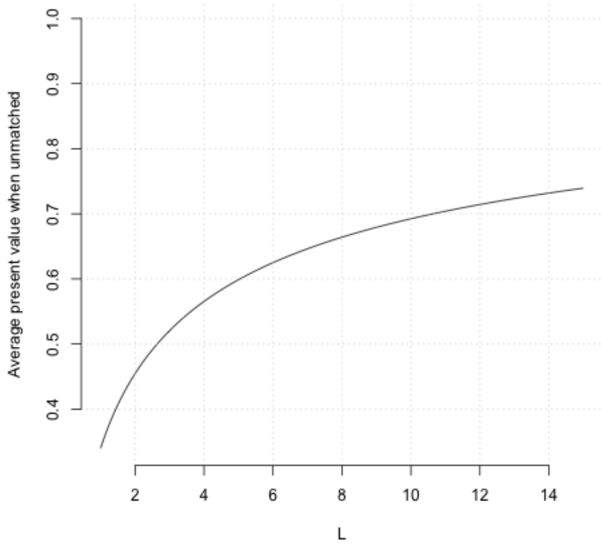
which equates inflows to and outflows from the population of unmatched users. Solving for u yields

$$u = \sqrt{\frac{\delta^2}{4(\eta^2\rho)^2} + \frac{\delta}{\eta^2\rho}L} - \frac{\delta}{2\eta^2\rho}.$$

We can substitute this expression for u into (19) to obtain an expression for the value of being unmatched as a function of the mass of users L on the platform and of other model primitives; this expression is increasing in L. Figure 7 plots the relationship between V_u and L under the choice of parameters $\rho = 0.5$, $\delta = 0.5$, $\mu_f = 1$, $\eta = 1$, and $r = 0.10$. The positive relationship between L and the consumer's value of joining the platform as an unmatched

[5] The average present value is defined as the net present value times the interest rate r.

Figure 7: Relationship between platform usage and value

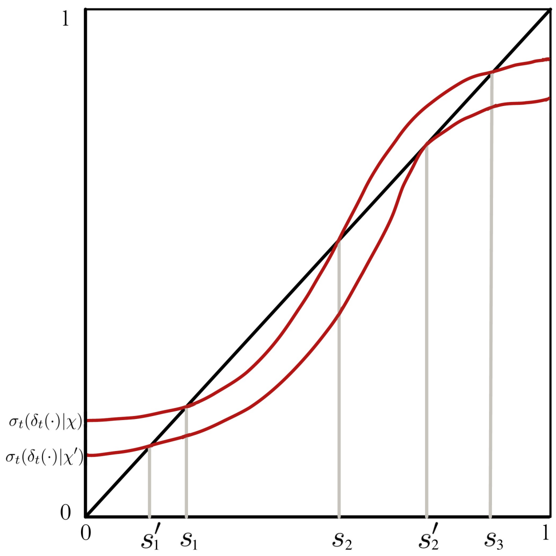


user justifies the inclusion of a network externality term in the indirect utilities of the model considered in the main text.

B Discussion of locally unique equilibria

This appendix discusses the concept of local uniqueness of equilibria as proposed in Section 3.1. Recall the condition that (4) must be nonsingular to ensure local uniqueness; I analyze this condition in the context of a binary logit model with positive network externalities. Figure 8 illustrates this model under two distinct values of the market characteristics χ_t. Equilibria in this model correspond to intersections between the s-shaped curve representing the mapping $s \mapsto \sigma_t(\delta_t(s)|\chi_t)$ and the 45-degree line. I interpret χ_t as the good's exogenous vertical quality; in the figure, decreasing χ_t from χ to χ' shifts the s-shaped curve downward until the equilibria s_2 and s_3 collapse into a single equilibrium s_2'. The matrix (4) is singular in this illustrative model when $\sigma_t(\delta_t(\cdot), \chi_t)$ is tangent to the 45-degree line, as happens when $\chi_t = \chi'$ at the equilibrium s_2'. We cannot define a unique equilibrium surface around (χ', s_2') because there are two distinct equilibrium market shares nearby s_2' when we marginally increase χ_t and no equilibrium market share near s_2' when we marginally decrease χ_t. This illustration suggests that local uniqueness will fail only in knife-edge cases.

Figure 8: Multiple equilibria in a binary choice model with network externalities



C Identification appendix

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

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

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

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

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

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

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

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

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

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

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

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

Proof. Inverting the equation

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

which is legal by Assumption INVERT-MARKET, yields

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

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

$$s_{jt} = \underbrace{\bar{\sigma}_j^{-1}(s_t) - g_j(x_{jt})}_{=:\kappa_j(s_t, x_{jt})} + \xi_{jt}. \tag{20}$$

Equation (20) falls within the NPIV framework: the left-hand side is observable and $\kappa_j(s, x_j)$ is an unknown function of (s, x_j). Consequently, the function κ_j is identified by the non-parametric instrumental variables argument of Newey and Powell (2003) under Assumption NPIV-MARKET. The identification of ξ_j follows immediately.

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

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

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

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

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

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

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

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

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

C.2 Proof of Proposition 2

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

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

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

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

for

$$\kappa_j(s, M) = \bar{\sigma}_j^{-1}(s) - h_j(Ms). \tag{21}$$

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

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

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

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

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

Now consider assumption (b). Note that

$$\frac{\partial \kappa_j}{\partial s}(s, M) = \nabla\bar{\sigma}_j^{-1}(s) - M\nabla h_j(Ms), \tag{22}$$

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

$$\lim_{M \downarrow 0} \frac{\partial \kappa_j}{\partial s}(s, M) = \nabla\bar{\sigma}_j^{-1}(s), \tag{23}$$

which shows how the large support assumption leads to identification of $\nabla\bar{\sigma}_j^{-1}$; the right-hand side of (23) is identified because (i) κ_j is identified on the full support of $\{s_t, M_t\}$ and (ii) the closure of this support includes $M_t = 0$ for all s_t. Note that the identification of ∇h_j consequently follows from (22). The levels of $\bar{\sigma}_j^{-1}$ and h_j are subsequently identified using the location normalization that $h_j(q_j^*) = 0$ for a known vector q_j^*. Indeed, the fundamental theorem of calculus for line integrals implies that for a differentiable path r with $r(0) = q_j^*$ and $r(1) = q_j$ in the support of $M_t s_t$,

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

Integration of ∇h_j over a path r in the support of $M_t s_t$ is justified by the convexity of this support as stipulated by Assumption CALC-TOT. Given the definition of κ_j in (21), the identification of h_j on the support of $M_t s_t$ and the previously established identification of κ_j on the support of $\{s_t, M_t\}$ implies the identification of $\bar{\sigma}^{-1}$ on the support of s_t. The identification of $\bar{\sigma}^{-1}$ on the support of s_t straightforwardly implies the identification of $\bar{\sigma}$ on the support of $\delta(x_t, s_t, \xi_t)$. □

C.3 Proof of Proposition 3

Before presenting the proof, I provide technical conditions that adapt Assumption 5 of Berry and Haile (2022) to my setting.

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

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

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

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

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

Proof. The proof closely follows Berry and Haile (2022). Let $w^*(\mathscr{s}, s, \xi)$ be the vector of individual characteristics that give rise to choice probabilities $\mathscr{s}$ when market shares equal s and the vector of unobservable product qualities is equal to ξ. Such a vector exists for all $\mathscr{s}$ in the support of $\mathscr{s}$ conditional on $s_{it} = s$ and $\xi = \xi_t$.. Additionally, $w^*(\mathscr{s}, s, \xi)$ is unique by virtue of Assumptions INV-DEMAND and INV-INDEX. Next, let $\mathcal{W} = \operatorname{supp} w_{it}$ and let $\|\cdot\|$ be the Euclidean norm. By Lemma 1 in Berry and Haile (2022), there exists a $s \in \operatorname{supp} s_t$ and $\Delta > 0$ such that for all w and w' in $\operatorname{supp} w_{it}$ such that $\|w - w'\| < \Delta$, there exist a choice probability vector $\mathscr{s}$ and vectors ξ and ξ' in $\operatorname{supp} \xi_t \mid s$ such that $w = w^*(\mathscr{s}, s, \xi)$ and $w' = w^*(\mathscr{s}, s, \xi')$. This follows from Lemma 1 after substituting s_t in place of their price vector P_t and recalling that I condition on exogenous market characteristics x_t throughout my analysis. Next, Lemma 2 in Berry and Haile (2022) implies that there is a $\Delta > 0$

such that for almost all w and w' in $\operatorname{supp} w_{it}$ such that $\|w - w'\| < \Delta$, $(Dg(w))^{-1}Dg(w')$ is identified. By Lemma 3 in Berry and Haile (2022), g is identified.

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

$$\begin{aligned} \mathscr{s}_{ijt} &= \sigma_j(\delta(w_{it}, \xi_t), s_t) \\ \delta_j(w_{it}, \xi_t) &= g_j(w_{it}) + \xi_{jt} \end{aligned} \tag{24}$$

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

$$g_j(w^*(\mathscr{s}^*, s_t, \xi_t)) = \sigma_j^{-1}(\mathscr{s}^*, s_t) - \xi_{jt} \tag{25}$$

We obtain this equation by σ with respect to its first argument in (24) at $w_{it} = w^*(\mathscr{s}, s_t, \xi_{it})$, where $\mathscr{s}^*$ is the common choice probability of Assumption CPROB; this inversion is justified by Assumption INV-DEMAND. The common choice probability $\mathscr{s}^*$ and the left-hand side of (25) are known. Therefore, (25) is a standard nonparametric regression equation with dependent variable $g_j(w^*(\mathscr{s}^*, s, \xi))$, nonparametric regression function $s \mapsto \sigma^{-1}(\mathscr{s}^*, s)$, and additive disturbance $-\xi_{jt}$. By the argument of Newey and Powell (2003), Assumptions NPIV-EX and NPIV-C identify each of the ξ_{jt} unobservables Given the choice probability function is

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

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

C.4 Identification of models with multiple demographic groups

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

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

$$\bar{\sigma}_{j,d}(x_t, S_t M_t^{\dagger}, \xi_t) = s_{jt}^d, \tag{26}$$

which is the analogue of (3) for the setting with multiple market shares. In (26),

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

so that

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

I study identification under the following index structure:

$$\begin{aligned} \bar{\sigma}_{j,d}(x_t, S_t M_t^{\dagger}, \xi_t) &= \bar{\sigma}_j^d(\delta_j^d(x_t, S_t M_t^{\dagger}, \xi_t)) \\ \delta_j^d(x_{jt}, S_t M_t^{\dagger}, \xi_t) &= x_{jt}^d + h_j^d(S_t M_t^{\dagger}) + \xi_{jt}^d \end{aligned} \tag{27}$$

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

The index structure (27) above restricts each x_{jt}^d to be scalar-valued, but the decomposition of x_{jt} into d-specific components is not necessarily an assumption; indeed, it remains possible to set $x_{jt}^d = x_{jt}$ for a single product characteristic x_{jt}. The assignment of a coefficient of one to x_{jt}^d is a scale normalization.

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

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

Re-arranging terms yields a nonparametric regression equation:

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

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

when group d consumers do not value $s_t^{d'}$.

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

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

$$\begin{aligned}\jmath_{ijt} &= \sigma_{j,d(i)}(\delta(w_{it},\xi_t),x_t,s_t) \\ \delta_j^d(w_{it},\xi_t) &= g_j^d(w_{it}) + \xi_{jt}^d\end{aligned} \tag{28}$$

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

Under suitably generalized assumptions, Proposition 3 is generalized to identify the model with multiple demographic groups. Indeed, the g_j^d functions are identified by applying the argument of Berry and Haile (2022) for identifying g as summarized in the proof of Proposition 3 (see Appendix C.3) to each demographic group separately. With the g_d^j functions in hand, one can proceed with identification using nonparametric regression equations of the form (25) that are specific to individual demographic groups d. As noted in Appendix C.3, identifying the ξ_t unobservables in these equations is sufficient for identification.

D Price sensitivity estimation

This section provides additional details of the price-sensitivity estimation procedure described in Section 5.4. To begin, I define $\bar{\sigma}_{jt}(p,s)$ as the mean probability that a consumer in market t uses site j under prices p when she believes that the prevailing market shares are s. The mean in the definition of $\bar{\sigma}_{jt}$ is taken over t's distribution of consumer characteristics w_{ijt} and unobservables ε_{ijt}. The market shares s need not be the market shares consistent with the mean choice probabilities $\bar{\sigma}_j(p,s)$; they are just a member of $\Delta^J = \{s \in (0,1)^J : \sum_{j=1}^J s_j \leq 1\}$. Let $\sigma_{jt}(p_t)$ denote the market shares that prevail under prices p_t. The function σ_t is implicitly defined by

$$\bar{\sigma}_t(p,\sigma_t(p)) = \sigma_t(p). \tag{29}$$

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

$$D_p\sigma_t(p) = [I - D_s\bar{\sigma}_t(p,\sigma_t(p))]^{-1}\, D_p\bar{\sigma}_t(p,\sigma_t(p)). \tag{30}$$

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

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

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

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

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

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

we have

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

Now note that

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

does not depend on α. This makes it convenient to write (30) 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 FOC(16) as

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

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

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

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

$$\alpha = -\frac{1}{J} \sum_j \frac{\sum_t M_t \sigma_{jt}(p_t)}{\sum_t M_t (\tilde{D}_p \sigma_t(p_t))_{jj} p_j}. \tag{31}$$

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 (31) yields my estimator of α:

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

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

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

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

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

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

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

The corresponding FOC is

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

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

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

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 price derivatives are $JD \times J$ and (ii) the specification above includes both share-type and quantity-type models since the f functions can implicitly depend on the populations of demographic groups within particular markets. I now provide the forms of the $D_s\bar{\sigma}$ and $D_p\bar{\sigma}$ functions under the multinomial logit parametric assumption:

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

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

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

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

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

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

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

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

D.1 Standard errors

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

E Additional results

Table 18 provides results of the microstep estimation for the baseline specification. Table 20 provides results of the microstep estimation for the "Age" demographic group specification. For each specification, standard errors are computed from the inverse Hessian estimate of the estimator's asymptotic variance.

Table 18: First-stage parameter estimates – "Overall" demographic group specification, demographic variables

	eharmony.com	match.com	okcupid.com	pof.com
Education: High school or less (Omit.)	0.000	0.000	0.000	0.000
	-	-	-	-
Education: Some college	0.072	0.106	-0.071	0.369
	(0.150)	(0.103)	(0.326)	(0.228)
Education: College degree	-0.076	-0.063	-0.293	-0.206
	(0.161)	(0.109)	(0.355)	(0.273)
Education: Advanced degree	-0.028	0.048	-1.329	-0.139
	(0.182)	(0.121)	(0.627)	(0.304)
Education: Unknown	0.230	0.020	-0.356	0.462
	(0.112)	(0.078)	(0.238)	(0.179)
Age: Under 25yo (Omit.)	0.000	0.000	0.000	0.000
	-	-	-	-
Age: 25-29yo	0.141	0.285	-0.110	0.344
	(0.178)	(0.135)	(0.411)	(0.253)
Age: 30-34yo	0.063	0.294	-0.460	0.424
	(0.167)	(0.126)	(0.402)	(0.238)
Age: 35-39yo	0.026	0.169	-0.539	0.266
	(0.163)	(0.124)	(0.382)	(0.236)
Age: 40-49yo	0.108	0.215	-0.216	0.013
	(0.154)	(0.118)	(0.343)	(0.226)
Age: 50+yo	0.097	0.249	-0.326	0.187
	(0.153)	(0.117)	(0.343)	(0.223)
Children in HH: No (Omit.)	0.000	0.000	0.000	0.000
	-	-	-	-
Children in HH: Yes	0.094	0.055	0.028	-0.045
	(0.067)	(0.050)	(0.180)	(0.097)
Race: White (Omit.)	0.000	0.000	0.000	0.000
	-	-	-	-
Race: Black	-0.082	-0.496	-0.587	-0.369
	(0.089)	(0.077)	(0.322)	(0.150)
Race: Asian	-0.287	-0.185	0.252	-0.993
	(0.216)	(0.142)	(0.388)	(0.510)
Race: Other	-0.293	-0.280	0.228	0.170
	(0.192)	(0.136)	(0.392)	(0.247)
Broadband: No (Omit.)	0.000	0.000	0.000	0.000
	-	-	-	-
Broadband: Yes	-0.496	-0.116	-0.131	-0.570
	(0.106)	(0.093)	(0.372)	(0.142)
Hispanic: No (Omit.)	0.000	0.000	0.000	0.000
	-	-	-	-
Hispanic: Yes	-0.107	0.069	-0.085	-0.208
	(0.059)	(0.043)	(0.168)	(0.091)
Income: Under 25k (Omit.)	0.000	0.000	0.000	0.000
	-	-	-	-
Income: 25-75k	0.089	-0.024	0.048	0.077
	(0.064)	(0.047)	(0.180)	(0.090)
Income: 75-100k	0.068	-0.054	0.061	-0.048
	(0.080)	(0.059)	(0.222)	(0.117)
Income: Over 100k	0.001	-0.052	0.306	-0.297
	(0.079)	(0.057)	(0.204)	(0.120)
HH size: 1 (Omit.)	0.000	0.000	0.000	0.000
	-	-	-	-
HH size: 2	-0.136	-0.206	-0.228	-0.008
	(0.098)	(0.072)	(0.255)	(0.146)
HH size: 3	-0.323	-0.295	-0.318	-0.037
	(0.117)	(0.086)	(0.306)	(0.172)
HH size: Over 3	-0.233	-0.357	-0.438	-0.056
	(0.114)	(0.084)	(0.300)	(0.169)
Log local population	0.042	-0.023	-0.097	-0.061
	(0.021)	(0.015)	(0.054)	(0.029)

Table 19: First-stage parameter estimates – "Overall" demographic group specification, web usage variables

	`eharmony.com`	`match.com`	`okcupid.com`	`pof.com`
Log Pages Viewed	0.141	0.237	0.548	0.100
	(0.073)	(0.057)	(0.201)	(0.104)
Log Browsing Duration	-0.136	-0.094	0.016	-0.022
	(0.060)	(0.044)	(0.154)	(0.085)
Pages Viewed: Adult	0.062	0.042	-0.022	0.078
	(0.012)	(0.011)	(0.050)	(0.012)
Pages Viewed: Advert	-0.011	-0.001	-0.001	-0.003
	(0.011)	(0.005)	(0.013)	(0.012)
Pages Viewed: Finance	0.030	0.017	-0.089	-0.126
	(0.027)	(0.019)	(0.084)	(0.050)
Pages Viewed: Gaming	-0.002	-0.033	-0.035	-0.012
	(0.011)	(0.011)	(0.032)	(0.020)
Pages Viewed: Government	0.018	0.010	-0.676	-0.108
	(0.028)	(0.021)	(0.346)	(0.088)
Pages Viewed: Info	-0.061	0.021	0.140	-0.087
	(0.053)	(0.030)	(0.046)	(0.087)
Pages Viewed: Malware	0.009	-0.017	-0.041	-0.013
	(0.007)	(0.006)	(0.023)	(0.014)
Pages Viewed: Media	-0.054	0.008	0.043	-0.174
	(0.021)	(0.011)	(0.021)	(0.044)
Pages Viewed: Other	0.000	-0.003	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Pages Viewed: Portal	0.016	0.033	0.011	0.024
	(0.004)	(0.003)	(0.009)	(0.005)
Pages Viewed: Retail	0.002	0.002	-0.000	0.004
	(0.004)	(0.003)	(0.010)	(0.006)
Pages Viewed: Social Media	-0.004	-0.003	-0.004	0.002
	(0.001)	(0.001)	(0.003)	(0.002)
Pages Viewed: Video	-0.021	-0.070	0.025	-0.076
	(0.012)	(0.011)	(0.013)	(0.025)
Pages Viewed: Weather	-0.102	-0.053	-0.135	-0.193
	(0.074)	(0.043)	(0.179)	(0.119)
Pages Viewed: Webservice	0.007	-0.004	0.010	-0.048
	(0.008)	(0.007)	(0.014)	(0.023)
Pages Viewed: Internet/Wireless	-0.006	-0.001	-0.021	-0.043
	(0.025)	(0.018)	(0.065)	(0.050)
Pages Viewed: News	-0.185	0.063	-0.026	0.047
	(0.076)	(0.031)	(0.123)	(0.064)
Pages Viewed: Sports	-0.123	0.011	-0.123	-0.043
	(0.049)	(0.022)	(0.103)	(0.056)
Pages Viewed: Travel	0.133	0.205	-0.992	-0.108
	(0.079)	(0.055)	(0.417)	(0.159)
Pages Viewed: Career	0.205	0.102	-0.127	0.166
	(0.062)	(0.054)	(0.275)	(0.095)
Pages Viewed: Downloads	0.110	-0.102	-0.388	0.160
	(0.108)	(0.109)	(0.495)	(0.146)
Pages Viewed: Directory	0.898	0.739	1.033	1.084
	(0.431)	(0.372)	(1.120)	(0.593)

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

	eharmony.com	match.com	okcupid.com	pof.com
Education: High school or less (Omit.)	0.000	0.000	0.000	0.000
	-	-	-	-
Education: Some college	0.071	0.112	-0.068	0.355
	(0.148)	(0.102)	(0.323)	(0.224)
Education: College degree	-0.078	-0.060	-0.295	-0.214
	(0.159)	(0.109)	(0.353)	(0.270)
Education: Advanced degree	-0.027	0.056	-1.313	-0.146
	(0.181)	(0.120)	(0.625)	(0.301)
Education: Unknown	0.228	0.024	-0.350	0.453
	(0.110)	(0.076)	(0.235)	(0.173)
Age: Under 25yo (Omit.)	0.000	0.000	0.000	0.000
	-	-	-	-
Age: 25-29yo	0.149	0.319	-0.094	0.357
	(0.171)	(0.131)	(0.407)	(0.244)
Age: 30-34yo	0.059	0.321	-0.451	0.428
	(0.161)	(0.122)	(0.397)	(0.228)
Age: 35-39yo	-2.006	-2.042	-7.552	-2.154
	(0.156)	(0.120)	(0.376)	(0.226)
Age: 40-49yo	-1.923	-1.995	-7.236	-2.409
	(0.146)	(0.113)	(0.335)	(0.214)
Age: 50+yo	-1.933	-1.962	-7.343	-2.231
	(0.145)	(0.112)	(0.334)	(0.210)
Children in HH: No (Omit.)	0.000	0.000	0.000	0.000
	-	-	-	-
Children in HH: Yes	0.093	0.056	0.017	-0.041
	(0.067)	(0.050)	(0.180)	(0.097)
Race: White (Omit.)	0.000	0.000	0.000	0.000
	-	-	-	-
Race: Black	-0.085	-0.497	-0.584	-0.369
	(0.087)	(0.076)	(0.318)	(0.148)
Race: Asian	-0.286	-0.187	0.281	-0.996
	(0.213)	(0.140)	(0.381)	(0.508)
Race: Other	-0.300	-0.282	0.248	0.172
	(0.189)	(0.135)	(0.383)	(0.242)
Broadband: No (Omit.)	0.000	0.000	0.000	0.000
	-	-	-	-
Broadband: Yes	-0.498	-0.115	-0.131	-0.572
	(0.106)	(0.093)	(0.368)	(0.142)
Hispanic: No (Omit.)	0.000	0.000	0.000	0.000
	-	-	-	-
Hispanic: Yes	-0.108	0.069	-0.084	-0.206
	(0.059)	(0.042)	(0.166)	(0.090)
Income: Under 25k (Omit.)	0.000	0.000	0.000	0.000
	-	-	-	-
Income: 25-75k	0.081	-0.026	0.052	0.082
	(0.064)	(0.047)	(0.178)	(0.090)
Income: 75-100k	0.063	-0.057	0.066	-0.046
	(0.079)	(0.059)	(0.221)	(0.116)
Income: Over 100k	-0.004	-0.053	0.316	-0.294
	(0.078)	(0.056)	(0.201)	(0.118)
HH size: 1 (Omit.)	0.000	0.000	0.000	0.000
	-	-	-	-
HH size: 2	-0.142	-0.212	-0.228	-0.014
	(0.097)	(0.072)	(0.252)	(0.143)
HH size: 3	-0.328	-0.300	-0.306	-0.040
	(0.116)	(0.085)	(0.303)	(0.170)
HH size: Over 3	-0.239	-0.362	-0.439	-0.064
	(0.113)	(0.084)	(0.298)	(0.167)
Log local population	0.043	-0.023	-0.095	-0.062
	(0.016)	(0.012)	(0.044)	(0.023)

Table 21: First-stage parameter estimates ("Age" demographic group specification), web usage variables

	eharmony.com	match.com	okcupid.com	pof.com
Log Pages Viewed	0.142	0.241	0.538	0.094
	(0.067)	(0.051)	(0.174)	(0.096)
Log Browsing Duration	-0.136	-0.097	0.014	-0.020
	(0.060)	(0.044)	(0.152)	(0.085)
Pages Viewed: Adult	0.063	0.042	-0.023	0.077
	(0.012)	(0.011)	(0.049)	(0.012)
Pages Viewed: Advert	-0.011	-0.001	-0.001	-0.003
	(0.011)	(0.006)	(0.013)	(0.012)
Pages Viewed: Finance	0.031	0.018	-0.093	-0.126
	(0.026)	(0.019)	(0.083)	(0.050)
Pages Viewed: Gaming	-0.002	-0.033	-0.034	-0.010
	(0.011)	(0.011)	(0.032)	(0.019)
Pages Viewed: Government	0.016	0.007	-0.742	-0.106
	(0.028)	(0.021)	(0.345)	(0.086)
Pages Viewed: Info	-0.057	0.021	0.140	-0.090
	(0.053)	(0.030)	(0.046)	(0.086)
Pages Viewed: Malware	0.009	-0.017	-0.042	-0.013
	(0.007)	(0.006)	(0.023)	(0.014)
Pages Viewed: Media	-0.055	0.008	0.042	-0.174
	(0.021)	(0.011)	(0.020)	(0.044)
Pages Viewed: Other	0.000	-0.004	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Pages Viewed: Portal	0.016	0.033	0.012	0.023
	(0.004)	(0.003)	(0.008)	(0.005)
Pages Viewed: Retail	0.002	0.002	0.000	0.005
	(0.004)	(0.003)	(0.010)	(0.006)
Pages Viewed: Social Media	-0.004	-0.003	-0.004	0.002
	(0.001)	(0.001)	(0.002)	(0.002)
Pages Viewed: Video	-0.022	-0.070	0.025	-0.076
	(0.012)	(0.011)	(0.012)	(0.025)
Pages Viewed: Weather	-0.105	-0.057	-0.130	-0.193
	(0.074)	(0.043)	(0.179)	(0.119)
Pages Viewed: Webservice	0.007	-0.004	0.011	-0.049
	(0.008)	(0.007)	(0.014)	(0.023)
Pages Viewed: Internet/Wireless	-0.006	-0.002	-0.025	-0.042
	(0.025)	(0.018)	(0.066)	(0.049)
Pages Viewed: News	-0.186	0.064	-0.037	0.052
	(0.076)	(0.031)	(0.124)	(0.064)
Pages Viewed: Sports	-0.123	0.014	-0.122	-0.041
	(0.049)	(0.022)	(0.103)	(0.056)
Pages Viewed: Travel	0.135	0.205	-0.998	-0.110
	(0.079)	(0.055)	(0.417)	(0.158)
Pages Viewed: Career	0.204	0.101	-0.123	0.168
	(0.062)	(0.054)	(0.272)	(0.094)
Pages Viewed: Downloads	0.106	-0.105	-0.387	0.151
	(0.107)	(0.109)	(0.492)	(0.146)
Pages Viewed: Directory	0.878	0.714	0.940	1.091
	(0.428)	(0.369)	(1.172)	(0.587)