University of Toronto
Department of Economics

Working Paper 758

Empirical Analysis of Network Effects in Nonlinear Pricing Data

By Liang Chen and Yao Luo

September 25, 2023

Empirical Analysis of Network Effects in Nonlinear Pricing Data*

Liang Chen† Yao Luo‡

September 25, 2023

Abstract

Network effects, i.e., an agent's utility may depend on other agents' choices, appear in many contracting situations. Empirically assessing them faces two challenges: an endogeneity problem in contract choice and a reflection problem in network effects. This paper proposes a nonparametric approach to tackle both challenges by exploiting restriction conditions from both demand and supply sides. We illustrate our methodology in the yellow pages advertising industry. Using advertising purchases and nonlinear price schedules from seven directories in Toronto, we find positive network effects, which account for a substantial portion of the publisher's profit and businesses' surpluses. We finally conduct counterfactuals to assess the overall and distributional welfare effects of the nonlinear pricing scheme relative to an alternative linear pricing scheme with and without network effects.

Key Words: Identification, Asymmetric Information, Network Effects, Nonlinear Pricing

JEL Codes: L11, L12, L13

*We thank Eugene Choo, Avi Goldfarb, Arvind Magesan, Isabelle Perrigne , and Quang Vuong for helpful comments. Luo thanks SSHRC for financial support. All remaining errors are our own.

†School of Economics, Zhejiang University, *Email: liang.chen@zju.edu.cn.*

‡Department of Economics, University of Toronto, *Email: yao.luo@utoronto.ca.*

1 Introduction

Nonlinear pricing has been analyzed extensively in various industries, particularly high-technology industries, such as broadband internet, cellular phone service, software applications, social media advertising, and among others. For instance, most broadband internet providers charge their services based on the transmission bandwidth and data usage; software manufacturers tie the prices of their products to the total expected usage of the software; emerging social media advertising is often charged according to display time and number of clicks; most cellular telephone service providers have extensive menus of usage-based pricing plans.

These industries also tend to have a salient feature, called network effects, as consumers often benefit from the products and/or services through social interactions. Network effects arise when the number of consumers using compatible products or services in a network influences the value that consumers derive from using the same network.[1] Cellular telephone service is a widely used example. The value of being part of a telecommunications network for cellular phone service users increases with the size of the network, as larger network size would also lead to increases in complementary services, which in turn benefits the cellular phone users. Other industries, such as broadband internet, computer softwares and social media advertising, share similar network effects.

While network effects have been shown to be important in many industries, they have been overlooked in the empirical nonlinear pricing literature. We bridge this gap by developing an empirical model that incorporates network effects into the seminal nonlinear pricing framework proposed by Maskin and Riley (1984). In our model, a monopoly firm designs a nonlinear pricing schedule to maximize its expected profit from selling products to heterogenous consumers. Consumers choose the quantities of products they purchase to maximize their utility. Network effects are modeled as consumer benefits from consuming products that depend on their expectations on the total purchases of their group members (group-level network effects) and all consumers in the market (market-level network effects).

The primary goal of this paper is to identify and estimate the consumer utility function and network effects from the observed data on purchases and nonlinear price schedules. They are key model primitives for policy analysis with important implications. Empiricists often confront two major problems when identifying these primitives. First, an endogeneity problem arises as the observed purchases of consumers are correlated with their unobserved characteristics, or inversely, as the observed price is corrected with unobserved product

[1] See Economides (1996) for a formal definition of network effects.

features. Second, there is a reflection problem to identify network effects in linear-in-means models, referred to as social interaction effects in the seminal paper by Manski (1993). That is, it is difficult to distinguish among exogenous effects (the influence of one's own observable characteristics), contextual effects (the influence of others' characteristics), and endogenous effects (the influence of others' behaviours), in linear-in-means models. It arises when the mean purchase in a group linearly depends on the group's characteristics.

The endogeneity problem has often been resolved using instrumental variables,[2] and the reflection problem has been addressed in a number of ways.[3] Luo, Perrigne and Vuong (2018) proposes a new approach to resolve the endogeneity problem in the nonlinear pricing framework of Maskin and Riley (1984). When credible instrumental variables are difficult to come by, their approach exploits equilibrium conditions from both the supply and demand sides of the market. This paper extends this idea to resolve the reflection problem. To the best of our knowledge, this is the first paper to apply Luo, Perrigne and Vuong (2018)'s idea of exploiting equilibrium conditions from both sides of the market for resolving the reflection problem and identifying network effects. We exploit the fact that consumers not only interact with their group members, but also interact with suppliers in the market.[4] The supplier's decisions, i.e., the seller's optimal pricing decisions, naturally endogenize the consumers' decisions. Previous literature on social interactions mainly focuses on the interactions between consumers, overlooking interactions between consumers and suppliers.

This paper proposes a new empirical approach to identify and estimate the consumer utility function and both group-level and market-level network effects from purchases and nonlinear pricing schedule data in multiple markets. Our approach mainly consists of three steps. First, we pick one market as the benchmark market and normalize the market-level network effects in that market to zero. We then express consumers' taste types and their intrinsic utility as functions of observed purchases and price tariffs from the benchmark market by exploiting the equilibrium conditions of both the consumers and the seller. These functions allow us to identify the distribution of consumer taste type and their intrinsic utility functions. Second, we decompose a consumer's taste type into individual-level exogenous effects, group-level contextual effects, and group-level endogenous network effects. Using the identified taste type distribution, we can identify the group-level network effects under

[2]See e.g., Berry, Levinsohn and Pakes (1995), Ekeland, Heckman and Nesheim (2004), and Heckman, Matzkin and Nesheim (2010).

[3]For example, in reduced-from regressions, Brock and Durlauf (2001b) introduces known nonlinearity or uses a dynamic analog, and Bramoullé, Djebbari and Fortin (2009) exploits social network structures to resolve the reflection problem.

[4]For instance, consumers are affected not only by other consumers in the market but also by the seller's price schedule.

a functional form specification. Third, we identify the consumers' taste type distribution function and market-level network effects for other markets. The consumer taste type distributions in other markets are identified following a similar procedure to step one. The market-level network effects in other markets are identified by exploiting the differences in price tariffs across markets.

This paper makes two main contributions. First, we provide a new empirical approach to identify and estimate a nonlinear pricing model featuring network effects. While network effects are important in many industries where nonlinear pricing is prevalent, network effects have been overlooked in the existing literature on structural analysis of nonlinear pricing. Second, we propose a new idea to resolve the reflection problem in nonlinear pricing contexts. Our idea highlights the importance of exploiting equilibrium conditions from both sides of the market. The existing literature solves the reflection problem using restrictions on social interaction structure, variations in group sizes, or network size. However, it is often difficult to obtain such information for nonlinear pricing data.

To illustrate our approach, we apply our approach to the yellow pages advertising industry, which is an earlier form of media platform exhibiting network effects. We consider a monopoly publisher sells advertising in yellow pages directories to local businesses. The monopoly publisher prices the advertising nonlinearly based on the size and quality of advertisements displayed in the directories. Businesses' benefit from purchasing advertising space depends not only on their own purchases, but also on their expected mean purchases in the same industry and mean purchases in the entire directory. These additional benefits represent network effects from two levels of interactions: industry level and market level.

Using purchases and price schedule data on yellow pages advertising from seven directories in Toronto,[5] we estimate the model primitives: businesses' marginal utility and network effects. We find positive network effects at both industry and market levels, which strengthens the findings in the seminal paper, Rysman (2004). Furthermore, businesses obtain a large amount of informational rents as the publisher screens heterogeneous businesses through nonlinear pricing. Informational rents account for approximately 90% of total revenues.

With the estimated model primitives at hand, we conduct counterfactual experiments to evaluate the welfare implications of the nonlinear pricing scheme. While nonlinear pricing increases seller profit, its welfare effects are distributional and often subject to policy debate. In particular, nonlinear pricing disproportionately impacts low-type buyers through exclusion and higher unit prices. See, e.g., Attanasio and Frayne (2006) for food pricing in Colombia,

[5] There is only one publisher, which produces and distributes all seven yellow pages directories in Toronto.

Borenstein (2012) for electricity pricing in California, Dalton (2014) for heath insurance pricing , and Luo, Perrigne and Vuong (2018) for mobile service pricing. As far as we know, our paper is the first to empirically quantify the distributional impacts of nonlinear pricing when network effects play important roles.

To do so, we conduct counterfactuals to simulate and compare the equilibrium outcomes under the nonlinear pricing scheme and an alternative linear pricing scheme with different levels of network effects. We find two main implications on the overall welfare from our counterfactual results. First, both the businesses and the publisher reap substantial benefits from network effects under both linear and nonlinear pricing schemes. Second, the publisher receives more revenue and businesses receive less surplus under the nonlinear pricing scheme than under the linear pricing scheme. This is consistent with the fact that a principal is able to extract more information rent from consumers under a nonlinear pricing scheme than under a linear pricing scheme. Moreover, our counterfactuals quantify the distributional impacts of nonlinear pricing when network effects play important roles. We find that nonlinear pricing affects businesses of different types disproportionately; businesses in the low-type group are most hurt by nonlinear pricing as these businesses pay higher unit prices. This result is consistent with the existing nonlinear pricing literature. We also find that network effects would attenuate the above distributional impacts of nonlinear pricing. This result is new to the literature. In sum, both network effects and pricing schemes significantly shape the overall and distributional welfare.

This paper is mainly related to the literature on structural analysis of nonlinear pricing. See, for instance, Bousquet and Ivaldi (1997), Miravete and Röller (2004), Huang (2008), Grubb (2009), Ascarza, Lambrecht and Vilcassim (2012), Luo, Perrigne and Vuong (2018), and Chen, Luo and Xiao (2019) on structural analysis of nonlinear pricing in cellular phone service markets, and Nevo, Turner and Williams (2016), Lambrecht, Seim and Skiera (2007), and Luo (2023) on structural analysis of nonlinear pricing in broadband internet service markets. Despite the importance of network effects, most of the existing literature on nonlinear pricing ignores network effects in the analysis.[6] In contrast, our paper extends the results in Luo, Perrigne and Vuong (2018) to account for important network effects. Moreover, while they analyze data from one market, we exploit variations in price schedules and purchases across markets.

This paper is also related to the literature on the identification of social interaction ef-

[6]Bousquet and Ivaldi (1997) is an exception. In that paper, authors parametrically estimate a demand model with consumption externalities under two-part tariff in telecommunications.

fects.[7] The seminal work, Manski (1993), addresses the reflection problem in linear-in-means models. Since then, a number of studies tackle this problem in various settings. For example, in linear models, Brock and Durlauf (2001b) and Bramoullé et al. (2009) show that the reflection problem can be solved by exploiting restrictions on the social interaction structure. Lee (2007), Moffitt (2001), and Graham (2008) achieve identification by using variations in group size.[8] In nonlinear models, the reflection problem may not exist and nonparametric identification may be achievable. See e.g.,Brock and Durlauf (2001a) and Brock and Durlauf (2007).[9] Finally, several recent important contributions find that the identification of social interaction effects can be achievable in settings with a large number of players. Xu (2018) establishes the identification of endogenous effects from a super-population large network, and Menzel (2016) shows that the interactions between players become negligible in the limit when the number of players goes to infinity. Our paper proposes a new approach to achieve identification of social interaction effects by exploiting the restriction conditions from both sides of the market.

Our empirical application is closely related to Rysman (2004) on analyzing the importance of network effects in the yellow pages industry. Our application distinguishes itself in several important ways. First, we hand-collected individual-level data on advertisement purchases, which allows us to model businesses as heterogeneous instead of homogeneous as in Rysman (2004). Second, instead of a quantity-setting model, we develop a nonlinear pricing model with network effects. Third, we show that our model is identified by simultaneously exploiting FOCs from both sides of the market. Rysman (2004) derives the FOCs of consumers, businesses, and publisher, and finds instrumental variables for each equation. Aryal and Gabrielli (2020) also studies this industry but ignores network effects. The authors focus on estimating the efficiency and distribution effects of competition between duopoly yellow pages print directories.

The remainder of this paper is organized as follows. Section 2 presents a nonlinear pricing model that features network effects at both industry and market levels. Section 3 provides a simple example to illustrate our identification idea and establishes the identification of model primitives. Section 4 develops a semi-parametric estimation procedure of the model

[7]See for Manski (2000), Durlauf and Ioannides (2010) and Blume, Brock, Durlauf and Ioannides (2011) surveys on this literature.

[8]Brock and Durlauf (2001b) and Bramoullé et al. (2009) show identification of network effects from known nonlinearity or dynamics of social interaction effects; Lee (2007) exploits variations in group size; Moffitt (2001) and Graham (2008) exploit variation in error term variance across groups to achieve identification of social interaction effects.

[9]Blume, Brock, Durlauf and Ioannides (2011) provides a comprehensive survey of social interactions in nonlinear models.

primitives. Section 5 runs an application to illustrate our empirical approach and conducts counterfactual experiments to investigate the importance of pricing schemes and network effects. Section 6 concludes. The Appendix collects all proofs, tables, and figures.

2 The Model

In this section, we present a general model, which extends the nonlinear pricing model developed by Maskin and Riley (1984) by incorporating network effects. To be consistent with our empirical application, the principal is a monopoly yellow pages directory publisher and the agents are businesses.[10] The publisher sells yellow pages advertisements to heterogeneous businesses who benefit from advertisements for reaching to local residents. Moreover, a salient feature in this industry is the publisher distributes yellow pages directories to residents for free.[11] The pricing scheme of advertisements is nonlinear and based on the quality and size displayed in the yellow pages directories.

2.1 Model Primitives

There are I businesses in a set $\mathcal{I} \equiv \{1, \cdots, I\}$, and each business belongs to one of J industries in a set $\mathcal{J} \equiv \{1, \cdots, J\}$. Businesses are heterogeneous and characterized by a scalar taste type for advertising $\theta \in [\underline{\theta}, \overline{\theta}]$ with $0 < \underline{\theta} < \overline{\theta} < \infty$. A business' taste type θ aggregates all his individual-level and industry-level characteristics that the publisher can not exploit through price discriminate. More specifically, we specify a nonparametric function, $\theta = \rho(x_i, x_j, \bar{q}_j, \nu_i)$, for business i belonging to industry j, where x_i is a vector of business i's observed individual characteristics capturing the exogenous effect; x_j is a vector of observed industry characteristics capturing the *contextual effects* – the influence of other businesses' characteristics in the same industry; $\bar{q}_j$ is the mean advertising quantity purchased by businesses in industry j capturing the *endogenous network effects* – the influence of purchase choices by other businesses in the same industry; ν_i is a component of the business' private characteristics known only by himself. Therefore, given a vector of mean purchases in each industry $\bar{\boldsymbol{q}} = (\bar{q}_1, \cdots, \bar{q}_J)$, the taste type is the private information of each business and the publisher knows only its distribution, $F(\cdot|\bar{\boldsymbol{q}})$, in the market. The market size of businesses in the market is normalized to one.

[10]The setup and notations can be adapted to any other market in which a monopoly (or approximate monopoly) sells products to heterogenous consumers using a nonlinear pricing scheme, for instance, cellular phone services markets studied in Luo, Perrigne and Vuong (2018).

[11]As shown in Halaburda and Yehezkel (2011), this so-called 'divide-and-conquer strategy' is a consequence of asymmetric information: the publisher finds it optimal to attract the side with the lower information problem – in this case, the residents.

Each business obtains utility $U(q, \theta, Q)$ from purchasing advertising and faces a nonlinear tariff $T(q)$, where q is the purchased quantity of advertising, and Q is the total purchased quantity of advertising in the market capturing the market-level network effects.[12] Recall that the taste type θ captures the contextual and endogenous network effects in each industry, which we hereafter refer to as industry-level network effects. We adopt a "macro" approach (see Economides (1996)) to model both industry-level and market-level network effects as their sources are not modeled explicitly. The publisher incurs a cost $C(Q)$ to produce a yellow pages directory with Q amount of advertising.

The model primitives are $[U(\cdot,\cdot,\cdot), \rho(\cdot,\cdot,\cdot,\cdot), F(\cdot|\bar{\boldsymbol{q}}), C(\cdot)]$. The next assumption follows the theoretical literature. A variable as a subscript indicates the derivative of a function with respect to this variable.

Assumption 1.

(i) *The utility function* $U(\cdot,\cdot,\cdot)$ *is twice continuously differentiable with* $U(\cdot,\cdot,\cdot) \geq 0$, $U_q(\cdot,\cdot,\cdot) > 0$, $U_{qq}(\cdot,\cdot,\cdot) < 0$, $U_{q\theta}(\cdot,\cdot,\cdot) > 0$ $U_{qq\theta}(\cdot,\cdot,\cdot) \geq 0$, *and* $U_{q\theta\theta}(\cdot,\cdot,\cdot) \leq 0$ *on* $[0,+\infty) \times [\underline{\theta}, \bar{\theta}] \times [0,+\infty)$.

(ii) *The aggregation function* $\rho(\cdot,\cdot,\cdot,\cdot)$ *is twice continuously differentiable on* $[-\infty,+\infty]^4$.

(iii) *The type distribution* $F(\cdot|\bar{\boldsymbol{q}})$ *is twice continuously differentiable with a density* $f(\cdot|\bar{\boldsymbol{q}}) > 0$ *on its support* $[\underline{\theta}, \bar{\theta}]$ *for any* $\bar{\boldsymbol{q}}$. *Moreover, the derivative of* $\theta - \{[1 - F(\theta|\bar{\boldsymbol{q}})/f(\theta|\bar{\boldsymbol{q}})]\}$ *is strictly positive on* $[\underline{\theta}, \bar{\theta}]$ *for any* $\bar{\boldsymbol{q}}$.

(iv) *The publisher's cost function* $C(\cdot)$ *is twice continuously differentiable on* $[0,+\infty)$ *with marginal cost* $C_Q(\cdot) > 0$, *and* $C_{QQ}(\cdot) \geq 0$ *on* $[0,+\infty)$.

Most parts in Assumption 1 are standard. Assumption 1-(**ii**) is specific to models with asymmetric information. The condition that the derivative of $\theta - \{[1 - F(\theta|\bar{\boldsymbol{q}})]/f(\theta|\bar{\boldsymbol{q}})]\}$ is strictly positive guarantees sorting in equilibrium, which is an important property required to solve the model.

2.2 The Information and Timing Structure

The model's information structure is specified as follows. Recall that business i's observed individual-level and industry-level characteristics, (x_i, x_j), are public information to

[12]The market-level network effects are interpreted as: the resident's usage of the yellow pages directory depends on the information provided captured by the total mount of advertising Q, which in turn determines the value of advertising for businesses.

all market participants, while his private characteristics ν_i are only known to himself. We assume that all market participants know the distribution of ν for all businesses in the market. Therefore, each business has private information on his own taste type θ; that is, each business knows its own taste type but does not know those of other businesses. The publisher knows only the distribution of the taste type in the market, $F(\cdot|\bar{\boldsymbol{q}})$, given any $\bar{\boldsymbol{q}}$. Since both the publisher and all businesses share the same relevant information, $F(\cdot|\bar{\boldsymbol{q}})$ is also known to all businesses in the market.

We focus on the interactions between the publisher (the monopoly supplier of advertising) and the businesses (the demanders of advertising). The time sequence of the interactions between the publisher and businesses is specified as follows:

(1) Given the public information, the publisher and all businesses formalize expectations on mean purchases $\bar{\boldsymbol{q}}^e = (\bar{q}_1^e, \cdots, \bar{q}_J^e)$ in all industries, and total purchase in the market, $Q^e = \sum_{j \in \mathcal{J}} \bar{q}_j^e$. We assume that the publisher and all businesses formalized the same levels of these expectations as they have the same information in the market.

(2) Given expectations $\bar{\boldsymbol{q}}^e$ and Q^e, the publisher designs a quantity-based price contract, $[q(\theta), T(q(\theta))]$, to maximize its expected profit.

(3) Given $\bar{\boldsymbol{q}}^e$ and Q^e, each business formalizes his own taste type θ, and chooses a quantity and price contract accordingly.

(4) Expectations, $\bar{\boldsymbol{q}}^e$ and Q^e are fulfilled in equilibrium.

2.3 The Optimization Conditions for Both Sides of the Market

On the demand side, each business formalizes a taste type θ given any expectations, $\bar{\boldsymbol{q}}^e$ and Q^e. A business of type θ chooses a quantity $q(\theta)$ as a function of his type and pays the corresponding tariff $T(q(\theta))$. This quantity maximizes the business' net payoff, $U(q, \theta, Q^e) - T(q)$ and solves the FOC,

$$T_q(q(\theta)) = U_q(q(\theta), \theta, Q^e), \tag{1}$$

which states that his marginal payoff must equal his marginal payment at quantity q. Equation (1) is the so-called incentive compatibility (IC) constraint. Since the payment of the standard listing is very small for businesses, we assume all businesses in the market always purchase advertising (full coverage) for simplicity.[13]

[13]The theoretical literature on nonlinear pricing usually makes this assumption in order to simplify the analysis or to make sharp predictions when the model is difficult to solve (see, e.g, Rochet and Stole (2002)

On the supply side, the publisher's expected profit, defined as its expected revenue minus the production cost is given by

$$\int_{\underline{\theta}}^{\overline{\theta}} T(q(\theta))f(\theta|\bar{\boldsymbol{q}})d\theta - C(Q^e) \tag{2}$$

for any expectations, $\bar{\boldsymbol{q}}^e$ and Q^e. The publisher chooses the optimal price contract, $[q(\cdot), T(\cdot)]$, subject to the IC constraint (1).[14] This gives us

$$\max_{q(\cdot),T(\cdot)} \int_{\underline{\theta}}^{\overline{\theta}} T(q(\theta))f(\theta|\bar{\boldsymbol{q}}^e)d\theta - C(Q^e), \tag{3}$$

subject to the IC constraint (1).

We assume that the solution that the publisher seeks is an optimal fulfilled expectation price contract in equilibrium. We solve it in two steps. First, we derive the publisher's optimal price contract for any given levels of expectations, $\bar{\boldsymbol{q}}^e$ and Q^e. Second, we find the fulfilled values of expectations in equilibrium, $\bar{\boldsymbol{q}}^*$ and Q^*, and the corresponding optimal price schedule. The following proposition provides the necessary conditions for the publisher's optimal fulfilled expectation price contract.

Proposition 1.

(i) *Under Assumption 1, the publisher's optimal fulfilled expectation price contract,* $[(q^*(\cdot), T^*(\cdot)]$ *satisfies the necessary conditions,*

$$U_q(q^*(\theta), \theta, Q^*) = C'(Q^*) + \frac{1 - F(\theta|\bar{\boldsymbol{q}}^*)}{f(\theta|\bar{\boldsymbol{q}}^*)} U_{q\theta}(q^*(\theta), \theta, Q^e), \tag{4}$$

$$T_q(q^*(\theta)) = U_q(q^*(\theta), \theta, Q^*) \tag{5}$$

for all $\theta \in [\underline{\theta}, \overline{\theta}]$*, where* $\bar{\boldsymbol{q}}^* = (\bar{q}_1^*, \cdots, \bar{q}_J^*)$ *and* Q^* *are the mean industry-level purchases and total market-level purchase fulfilled in equilibrium, respectively.*

(ii) *If* $U_q(q, \theta, Q)$ *is bounded for all* q, θ, Q*, then the optimal fulfilled expectation price contract always exists.*

Proof. See the Appendix. □

and Armstrong and Vickers (2001) which study nonlinear pricing models with competition.)

[14]Note that the publisher's optimization problem is not subject to the individual rationality constraint as we assume the coverage.

Equations (4) and (5) characterize the optimal pricing contract, $[q^*(\cdot), T^*(\cdot)]$, in equilibrium. Equation (4) characterizes the optimal condition for the supply side, which states that in equilibrium the publisher chooses a pricing contract ensuring the total marginal payoff of each business type equals the marginal cost plus a nonnegative distortion term (informational rent) arising from incomplete information. Equation (5), which characterizes the optimal condition for the demand side, says that at equilibrium each business type chooses a purchase ensuring the marginal payment equals the marginal payoff of that type. This condition implies that the mean equilibrium purchase in each industry depends on the tariff function.

3 Identification

This section establishes identification of model primitives from observables in multiple markets. The model primitives, $[U(\cdot,\cdot,\cdot), \rho(\cdot,\cdot,\cdot,\cdot), F^m(\cdot|\bar{\boldsymbol{q}}^m), C(\cdot)]$, are the business' utility function, the taste type aggregation function capturing industry-level network effects, the taste type distribution in market $m \in \{1, \cdots, M\}$ with an observed mean purchase $\bar{\boldsymbol{q}}^m$ for each market, the publisher's cost function, where M denotes the number of markets we observe in the data. The analyst observes all businesses' purchased quantities, price tariffs, and individual-level and industry-level characteristics from multiple markets. Thus, the identification investigates whether these model primitives can be uniquely recovered from the observables $[G^m(\cdot), T^m(\cdot), \mathcal{H}^m(\cdot,\cdot), Q^m]$, where $G^m(\cdot)$ is the purchase quantity distribution, $T^m(\cdot)$ is the tariff function, $\mathcal{H}^m(\cdot,\cdot)$ is the joint distribution of businesses' individual-level and industry-level characteristics, and Q^m is the total advertisement quantity of market m. Note that, here, we implicitly assume that, among all model primitives, only the taste type distribution varies across markets, which drivers the differences in observables across markets.

3.1 Discussion

The identification is related to the literature on identifying network effects in utility function $U(q, \theta, Q)$. Regarding the identification of market-level network effects embedded in market-level advertisement quantity Q, Brock and Durlauf (2001b) and Sweeting (2009) show that variations in the realized equilibria across markets provide the leverage required to identify models with social interactions. We follow their idea to identify the market-level network effects from data on multiple markets. Regarding the identification of industry-level network effects embedded in $\theta = \rho(\cdot,\cdot,\cdot,\cdot)$, Manski (1993) shows that contextual and endogenous effects are not identified in linear-in-means models due to the reflection problem arising from the endogeneity of mean purchases. The existing literature solves this problem from further restrictions on social interaction structure (Brock and Durlauf (2001b), Bramoullé et al. (2009)), variations in group size (Lee (2007), Moffitt (2001), Graham (2008)), or the size of networks (Xu (2018)). However, it is often difficult to obtain information on social interaction structure in nonlinear pricing data. We show that the FOCs for both sides of the market exploited in Luo, Perrigne and Vuong (2018) can also help to solve the reflection problem without further restrictions on social network structure or variations in group sizes.

In the remainder of this section, we provide a simple illustrative example to show how exploiting FOCs for both sides of the market helps to solve the reflection problem in nonlinear pricing models. We then provide a four-step nonparametric procedure to identify the model primitives.

3.2 An Illustrative Example

This section constructs a simple example to illustrate how FOCs for both sides of the market help to solve the reflection problem for identifying the industry-level network effects. To focus on our idea, we make the following parametric restrictions on the model primitives. The utility function is specified as

$$U(q_i, \theta_i) = \theta_i q_i - \frac{1}{2} q_i^2 \tag{6}$$

for any business $i \in \mathcal{I}$ with taste type θ_i in industry $j \in \mathcal{J}$ purchasing q_i units of advertising. It is strictly concave in q_i for all $\theta_i > q_i$ and increasing in θ_i for any positive q_i. The aggregation function of the taste type, $\rho(\cdot, \cdot, \cdot, \cdot)$, is assumed to be linear,

$$\theta_i = \alpha_0 + \alpha_1 x_i + \alpha_2 x_j + \alpha_3 \bar{q}_j + \nu_i, \tag{7}$$

where we assume x_i; $x_j = \frac{1}{n_j} \sum_{i \in j} x_i$ with n_j denoting the number of businesses in industry j.[15] Moreover, the publisher's cost function $C(Q)$ is also assumed to take the form, $C(Q) = Q$, for simplicity.

Under these restrictions, the necessary conditions (4) and (5) for the equilibrium optimal

[15]This setup of group-level characteristics is consistent with Manski (1993).

fulfilled expectation pricing contract in Proposition 1 are reduced to,

$$\theta_i - q_i = 1 + \frac{1 - F(\theta_i|\bar{\boldsymbol{q}})}{f(\theta_i|\bar{\boldsymbol{q}})}, \tag{8}$$

$$T_q(q_i) = \theta_i - q_i \tag{9}$$

for business i with taste type θ_i. Condition (8) is the publisher's FOC (on the supply side) and condition (9) is the business's FOC (on the demand side).

In the literature on social interactions, researchers are interested in recovering the parameters $(\alpha_0, \alpha_1, \alpha_2, \alpha_3)$ from the data on individuals in the network by only exploiting their FOCs of optimization, i.e., the data, $(x_i, x_j, q_i)_{i \in \mathcal{I}, j \in \mathcal{J}}$ and the FOC (9) of businesses in our context. When only exploiting the FOCs of individuals in the network, there is no micro-foundation to rationalize the cost function (tariff function $T(q_i)$). Researchers often assume that the cost function takes a quadratic form, which delivers simple linear regression models (see e.g., Blume et al. (2015)). Following this idea, we assume a quadratic tariff function, $T(q_i) = \frac{1}{2}q_i^2$,[16] in our example. Substituting it into FOC (9) and expanding θ_i, we obtain a regression model,

$$q_i = \frac{\alpha_0}{2} + \frac{\alpha_1}{2}x_i + \frac{\alpha_2}{2}x_j + \frac{\alpha_3}{2}\bar{q}_j + \frac{1}{2}\nu_i. \tag{10}$$

The regression model (10) is a classic linear-in-means model, first studied in the seminal paper by Manski (1993). It is well known that it suffers the reflection problem and is not identified. The reflection problem arises from the fact that the mean equilibrium purchase $\bar{q}_j$ linearly depends on x_j. To see this, taking expectations on both sides of (10) conditional on (x_i, x_j) and assuming $E[\nu_i|x_i, x_j] = 0$, we obtain the mean equilibrium purchase, $\bar{q}_j = \frac{\alpha_0 + (\alpha_1 + \alpha_2)x_j}{2 - \alpha_3}$, which shows linear dependence between $\bar{q}_j$ and x_j. Substituting it into (10) yields,

$$q_i = \frac{\alpha_0}{2 - \alpha_3} + \frac{\alpha_1}{2}x_i + \frac{2\alpha_2 + \alpha_1\alpha_3}{2(2 - \alpha_3)}x_j + \frac{1}{2}\nu_i, \tag{11}$$

which is not identified as we have only three independent variables $(1, x_i, x_j)$ to estimate four parameters $(\alpha_0, \alpha_1, \alpha_2, \alpha_3)$.

Manski (2000) pointed out that if $\bar{q}_j$ depends on x_j nonlinearly, the reflection problem

[16] Note that there is no loss of generality to assume the coefficient being $\frac{1}{2}$ in the tariff function. Alternatively, our main arguments still hold by simply assuming the tariff to be $T(q_i) = \gamma q_i^2$ for some observed coefficient $\gamma > 0$.

does not exist and the social interaction effects are identified. We now show that, when researchers also exploit the supply-side FOC (of the publisher), the nonlinear relationship between $\bar{q}_j$ and x_j is generically achieved. Recall the publisher's FOC,

$$\theta_i - q_i = 1 + \frac{1 - F(\theta_i|\bar{\boldsymbol{q}})}{f(\theta_i|\bar{\boldsymbol{q}})}, \tag{12}$$

which provides a micro-foundation to rationalize the marginal tariff function, $T_q(q_i) = \theta_i - q_i = 1 + \frac{1-F(\theta_i|\bar{\boldsymbol{q}})}{f(\theta_i|\bar{\boldsymbol{q}})}$.

Rearranging (12) yields,

$$q_i = \theta_i - 1 - \frac{1 - F(\theta_i|\bar{\boldsymbol{q}})}{f(\theta_i|\bar{\boldsymbol{q}})} \equiv H(\theta_i). \tag{13}$$

It can be shown that $H(\cdot)$ is non-linear in θ as long as the taste type distribution $f(\theta)$ is not an exponential distribution or a power law distribution.[17] Hence, the model parameters are identified. To see it, taking expectations on both sides of (13) conditional on (x_i, x_j), we obtain $\bar{q}_j = E_\nu[H(\theta_i)|x_i, x_j] = E_\nu[H(\alpha_0+\alpha_1 x_i+\alpha_2 x_j+\alpha_3\bar{q}_j+\nu_i)|x_i, x_j]$, which is a fixed-point mapping of q_j. Since $H(\cdot)$ is non-linear, it defines a nonlinear relationship between $\bar{q}_j$ and x_j, which is denoted by $\bar{q}_j = S(x_j; \alpha_0, \alpha_1, \alpha_2, \alpha_3)$. Substituting this nonlinear relationship into regression model (10) yields,

$$q_i = \frac{\alpha_0}{2} + \frac{\alpha_1}{2}x_i + \frac{\alpha_2}{2}x_j + \frac{\alpha_3}{2}S(x_j; \alpha_0, \alpha_1, \alpha_2, \alpha_3) + \frac{1}{2}\nu_i. \tag{14}$$

The parameters $(\alpha_0, \alpha_1, \alpha_2, \alpha_3)$ in this nonlinear regression model are obviously identified.

3.3 Identification Procedure

We propose a four-step nonparametric procedure to identify the model primitives, which exploits the FOCs of both the publisher and the businesses in the two-sided market. The first step identifies the marginal cost for the production function, $C(\cdot)$. The second step follows Luo, Perrigne and Vuong (2018) and identifies the utility function $U(\cdot,\cdot,\cdot)$ from the data on a picked benchmark market in which we normalize the payoff from market-level network effects to be zero. The taste type distribution in this benchmark market is also

[17] To see this, suppose first that $H(\theta_i)$ has the linear form, $\theta_i - 1 - \frac{1-F(\theta_i)}{f(\theta_i)} = \kappa\theta_i + \gamma$ for $\kappa \neq 1$. We can obtain the taste distribution function as $F(\theta_i) = 1 + c \cdot ((\kappa-1)\theta_i + \gamma + 1)^{\frac{1}{\kappa-1}}$, which is a power law distribution. Suppose then that $H(\theta_i)$ has the linear form, $\theta_i - 1 - \frac{1-F(\theta_i)}{f(\theta_i)} = \theta_i + \gamma$. We can obtain the taste distribution function as $F(\theta_i) = 1 + c \cdot e^{-\frac{1}{\gamma+1}\theta_i}$, which is an exponential distribution.

identified in this step. Once the taste type distribution is identified, we identify the taste type aggregation function $\rho(\cdot,\cdot,\cdot,\cdot)$ from the data in the benchmark market in the third step. The fourth step identifies the market-level network effects and taste type distribution for other markets.

The identification results are based on the following identifying assumptions.

Assumption 2.

(i) *The utility function $U(\cdot,\cdot,\cdot)$ is assumed to be multiplicatively separable in taste type θ with the form,*

$$U(q,\theta,Q) = \theta\Big[V(q) + W(q,Q)\Big], \tag{15}$$

and $V(0) = 0$, $W(0,Q) = 0$ for any Q.

(ii) *The payoff from market-level network effects is normalized to be zero in the benchmark market, i.e., $W(\cdot, Q^{\dagger}) = 0$, where $\dagger$ indexes the benchmark market.*

Assumption 2-(i) follows Luo, Perrigne and Vuong (2018) and assumes utility function $U(\cdot,\cdot,\cdot)$ to be multiplicatively separable in taste type θ. Furthermore, $V(\cdot)$ and $W(\cdot,\cdot)$ can be interpreted as the base intrinsic payoff from purchase and the payoff from market-level network effects, respectively, which are assumed to be additively separable. $V(0) = 0$ and $W(0,Q) = 0$ for any Q imply that businesses do not have any intrinsic payoff or any payoff from market-level network effects if they do not purchase advertising. Assumption 2-(ii) is a normalization. This normalization is necessary because utility is ordinal. We can only identify relative payoffs from market-level network effects across markets with different total purchases. Our identification exploits the necessary conditions (4) and (5) for both the publisher and businesses shown in Proposition 1 , which, under Assumption 2, can be written as

$$\theta\Big[V_q(q) + W_q(q,Q^m)\Big] = C_Q(Q^m) + \frac{1-F^m(\theta|\bar{\boldsymbol{q}}^m)}{f^m(\theta|\bar{\boldsymbol{q}}^m)}\Big[V_q(q) + W_q(q,Q^m)\Big], \tag{16}$$

$$T^m_q(q) = \theta\Big[V_q(q) + W_q(q,Q^m)\Big] \tag{17}$$

in market m, where $q = q^m(\theta)$ is the observed equilibrium purchase for each type-θ business in market m.

STEP 1: IDENTIFICATION OF $C(\cdot)$

Our first identification result concerns the publisher's cost function using data on price schedules. Multi-market data provide us with variations in the total purchase quantity Q. Since we observe both the price schedule $T(\cdot)$ and maximum purchase $\overline{q}$ for multiple markets, we can recover the marginal cost function $C_Q(Q)$ by exploiting variations in Q and $T(\cdot)$. Proposition 2 shows that only the marginal cost function is identified from observables.

Proposition 2. *The marginal cost function $C_Q(\cdot)$ is identified on $[\underline{Q}, \overline{Q}]$ where $\underline{Q}$ and $\overline{Q}$ denote the minimum and maximum total quantity purchased among all the markets, respectively.*

This result is not surprising since the model involves the cost function only through the marginal cost. To see this, substituting (16) into (17) and evaluating at $\theta = \overline{\theta}$, we obtain

$$T_q^m(\overline{q}^m) = C_Q(Q^m)$$

for any $Q^m \in [\underline{Q}, \overline{Q}]$ in market m. Multi-market data provide us with variations in both $T(\cdot)$ and Q, which allow us to identify the marginal cost function $C_Q(\cdot)$.[18]

Note that we identify only the marginal cost function $C_Q(\cdot)$. It is conceivable that there is a fixed cost to produce a yellow page directory. To estimate the fixed cost, we could use data on minimum total purchases $\underline{Q}$. As any firm would only operate if net profit is positive, we can reasonably believe that the publisher's profit and cost are equal. With estimated model primitives, we can simulate the profit level and variable cost in this market. We can then learn about the fixed cost, thereby identifying the cost function $C(\cdot)$.

STEP 2: IDENTIFICATION OF $V(\cdot)$ AND $F^{\dagger}(\cdot)$

In Step 2, we identify the base intrinsic payoff function and the taste type distribution function in the benchmark market. To do so, we exploit the necessary conditions (16) and (17) for the benchmark market, which is denoted by superscript † hereafter. The following lemma shows that businesses' marginal base intrinsic payoff $V_q(\cdot)$ and their unobserved type function $\theta^{\dagger}(\cdot)$ for the benchmark market can be expressed as functions of the observed quantity purchased q, its distribution $G^{\dagger}(q)$, and the observed distribution of pricing tariff $T^{\dagger}(\cdot)$ in the benchmark market.

Lemma 1. *The necessary conditions* (16) *and* (17) *in the benchmark market are equivalent*

[18] Although the institutional constraint on the actual maximum size of advertisement is double-page spread in the yellow pages industry, the quality-adjusted maximum size of advertisement can vary across markets.

to

$$V_q(q) = \frac{T_q^\dagger(q)}{\overline{\theta}}\xi^\dagger(q) \quad \text{and} \quad \theta^\dagger(q) = \frac{\overline{\theta}}{\xi^\dagger(q)} \tag{18}$$

for all $q \in (0, \overline{q}]$, *where*

$$\xi^\dagger(q) = \Big[1 - G^\dagger(q)\Big]^{1-\frac{C_{Q'}(Q^\dagger)}{T_q^\dagger(q)}} \exp\Big\{C_Q(Q^\dagger)\int_q^{\overline{q}} \frac{T_{zz}^\dagger(z)}{T_z^\dagger(z)^2}\log\Big[1 - G^\dagger(z)\Big]dz\Big\}, \tag{19}$$

with $\xi^\dagger(0) = \lim_{q\to 0}\xi^\dagger(q) = \overline{\theta}/\theta^\dagger(0)$ *and* $\xi^\dagger(\overline{q}) = 1$.

Proof. See the Appendix. □

Lemma 1 follows Luo, Perrigne and Vuong (2018). It exploits the one-to-one mapping between the taste type θ and purchases q. By Lemma 1, a natural normalization is $\overline{\theta} = 1$. The marginal base intrinsic payoff function $V_q(\cdot)$ can then be uniquely recovered on $(0, \overline{q}]$ from the observables in the benchmark market. Moreover, the taste type distribution in the benchmark market $F^\dagger(\cdot)$ can be uniquely recovered on $[\underline{\theta}, \overline{\theta}]$ from the same observables. The following assumption and proposition formalize this result.

Assumption 3. $\overline{\theta} = 1$.

Using such a normalization and recalling that $V(0) = 0$, we can identify the intrinsic payoff function $V(\cdot)$ as

$$V(q) = \int_0^q T_z^\dagger(z)\xi^\dagger(z)dz,$$

from (18).

Proposition 3. *Under Assumption 3, the base intrinsic payoff function* $V(\cdot)$ *and the business' taste type distribution* $F^\dagger(\cdot)$ *in the benchmark market are identified on* $(0, \overline{q}]$ *and* $[\underline{\theta}, \overline{\theta}]$, *respectively.*

STEP 3: IDENTIFICATION OF TASTE TYPE AGGREGATION FUNCTION $\rho(\cdot,\cdot,\cdot,\cdot)$

We now turn to the identification of the aggregation function $\theta = \rho(x_i, x_j, \bar{q}_j, \nu_i)$. Recall that x_i and x_j are observed individual-level and industry-level characteristics for business i in industry j, $\bar{q}_j$ is the mean business advertisement quantity purchased in industry j, and ν_i is the unobserved characteristics. Our identification consists of two steps using the data from the benchmark market. First, we identify taste type as a nonlinear function of purchase quantity from the benchmark market, $\theta^\dagger = \theta^\dagger(q)$, following Lemma 1. In particular,

$\theta^\dagger = 1/\xi^\dagger(q)$ from Equation 18 and Assumption 3. Second, with the estimated $\theta^\dagger$ in hand, the identification of $\rho(\cdot,\cdot,\cdot,\cdot)$ can be achieved under two conditions: that $\theta^\dagger(q)$ is nonlinear, and the mean of ν is zero. We consider a linear aggregator of $\theta^\dagger$ as in the illustrative example, $\theta_i^\dagger = \alpha_0 + \alpha_1 x_i^\dagger + \alpha_2 x_j^\dagger + \alpha_3 \bar{q}_j^\dagger + \nu_i$, where $\bar{q}_j^\dagger = \frac{1}{n_j}\sum_{i\in j} q_i^\dagger$ is the mean purchase of businesses in industry j. Inverting $\theta^\dagger = 1/\xi^\dagger(q)$ for all i, we have $q_i^\dagger = \xi^{\dagger-1}(1/\theta_i^\dagger)$, where $\xi^\dagger(\cdot)$ is identified from Step 2. Substituting it into the above specification of $\theta_i^\dagger$ yields a new regression equation $\theta_i^\dagger = \alpha_0 + \alpha_1 x_i^\dagger + \alpha_2 x_j^\dagger + \alpha_3 \frac{1}{n_j}\sum_{i\in j} \xi^{\dagger-1}(1/\theta_i) + \nu_i$. As $\xi^{\dagger-1}(\cdot)$ is a known nonlinear function, the parameters in the specification are identified as illustrated in our example.

STEP 4: IDENTIFICATION OF $W(\cdot,\cdot)$ AND $F^m(\cdot)$

Finally, we show the identification of the payoff function from market-level network effects $W(\cdot,\cdot)$ and businesses' taste type distributions $F^m(\cdot)$ in any other market m. Similar to the identification in the benchmark market, we can identify the sum of marginal intrinsic payoff functions and marginal payoff functions from market-level network effects. Applying (18) in Lemma 1 to other markets gives $V_q(\cdot) + W_q(\cdot, Q_m) = T_q^m(\cdot)\xi^m(\cdot)$, where $\xi^m(\cdot)$ is obtained by replacing "†" with "m" in (19) . Moreover, the taste type distribution in this market is also identified, as the mapping from purchase quantity to taste type can be written in terms of observables; that is, $\theta^m(\cdot) = 1/\xi^m(\cdot)$. Since $V_q(\cdot)$ is identified from $V_q(\cdot) = T_q^\dagger(\cdot)\xi^\dagger(\cdot)$ using the data from the benchmark market, the marginal payoff function from market-level network effects can be identified from $W_q(\cdot, Q_m) = T_q^m(\cdot)\xi^m(\cdot) - T_q^\dagger(\cdot)\xi^\dagger(\cdot)$. Recalling that $W(0,Q) = 0$ for any Q, we can identify the payoff function from market-level network effects $W(\cdot, Q^m)$ in market m using

$$W(q,Q) = \int_0^q \Big[T_z^m(z)\xi^m(z) - T_z^\dagger(z)\xi^\dagger(z)\Big]dz. \tag{20}$$

The following proposition formalizes these results.

Proposition 4. *The network payoff $W(\cdot,\cdot)$ and the business' taste type distribution in market m, $F^m(\cdot)$, are identified on $(0,\overline{q}] \times [\underline{Q},\overline{Q}]$ and $[\underline{\theta},\overline{\theta}]$, respectively.*

Proposition 4 shows that the payoff function from market-level network effects, $W(\cdot, Q^m)$, and businesses' taste type θ^m in any market m can be uniquely recovered on $(0,\overline{q}]$ from the observables $[G^m(\cdot),\ T_q^m(\cdot)]$ and the identified $V_q(\cdot)$. Intuitively, the identification of the payoff function from market-level network effects is achieved by using the variations of price schedule $T(\cdot)$ and advertising purchase distribution $G(\cdot)$ across markets.

4 Estimation

In view of the identification results, we know that the model primitives are identified from tariff functions, purchase distributions, and distributions of businesses' individual-level and industry-level characteristics, $[T^m(\cdot), G^m(\cdot), \mathcal{H}^m(\cdot,\cdot)]$ for market m. Following the identification results, we propose a four-step procedure to estimate the model primitives using data from multiple markets. In the first step, we first construct tariff function $T^m(q)$ using the observed price schedule for any market m. The marginal cost function $C_Q(\cdot)$ is then estimated from $C_Q(Q^m) = T_q^m(\overline{q}^m)$ following Proposition 2. In the second step, we first estimate the purchase distribution $G^m(\cdot)$ from purchase data in any market m . We then pick a benchmark market to estimate the marginal intrinsic payoff function $V_q(\cdot)$ and businesses' pseudo taste type $\theta^m(\cdot)$ for any market m following Lemma 1. In the third step, we estimate the payoff function from market-level network effects for other markets, $W(\cdot, Q^m)$ where $m \neq \dagger$, using (20). In the fourth step, we estimate the type aggregation function $\rho(\cdot,\cdot,\cdot,\cdot)$ nonparametrically from businesses' estimated pseudo taste types of businesses and observed individual-level and industry-level characteristics. We also estimate the businesses' taste type density distribution functions $f^m(\cdot)$ nonparametrically from the estimated pseudo taste types of businesses in any market m. We discuss these steps in detail below.

First, we construct tariff function $T^m(q)$ using the observed price schedule for any market m. To be consistent with our empirical application, we illustrate it in the context of yellow page advertisements. It is well known that yellow page advertisements differ in their display size and quality (such as color and other features). For this reason, we follow Busse and Rysman (2005) to fit the shape of the price schedules and construct a quality-adjusted quantity index which aggregates size, color, and other features of advertisements.[19] Moreover, we exploit the price schedule for multicolored displays and adjust the sizes for other colors accordingly.

In particular, using the price schedules for multicolored display advertisements from multiple markets, we estimate the following equation:

$$\log T_j^m = \beta_0 + \beta_1 \log S_j^m + \beta_2 \log Q^m + \epsilon_j^m, \tag{21}$$

where T_j^m is the price in dollors for multicolored display j in market m, S_j^m is the advertising size measured in square picas, and Q^m is the total quantity purchases in market m measured

[19]It may be useful to integrate quality in the nonlinear pricing model. This may require a complex multidimensional screening model, which is known to be difficult to solve. See Armstrong (1996) and Aryal and Gabrielli (2020).

in square picas. We then use the regression result to construct the tariff functions and quality-adjusted quantities for other color advertisements. More specifically, the quality-adjusted quantity for any observed purchases in market m is constructed by

$$q = e^{-\hat{\beta}_0/\hat{\beta}_1} t^{1/\hat{\beta}_1} (Q^m)^{-\hat{\beta}_2/\hat{\beta}_1}, \tag{22}$$

where $\hat{\beta}_0, \hat{\beta}_1$ and $\hat{\beta}_2$ are estimated coefficients from (21), q is quality-adjusted quantity, and t is observed price for any business type advertisement.

The tariff function for market m with total quantity Q^m is constructed by

$$\hat{T}^m(q) = e^{\hat{\beta}_0} q^{\hat{\beta}_1} (Q^m)^{\hat{\beta}_2}, \tag{23}$$

where $\hat{T}^m(q)$ is the constructed tariff function for the market with total quantity Q^m and q is the quality-adjusted quantity. Using (23), the marginal cost for market m can be estimated by $\hat{C}_Q(Q^m) = \hat{T}^m_q(\overline{q}^m)$, following which the marginal cost function can be estimated accordingly.

Second, we estimate the purchase distribution $G^m(\cdot)$ and businesses' pseudo taste type $\theta^m(\cdot)$ for any market m, and the marginal intrinsic payoff function $V_q(\cdot)$. Moreover, the business taste type function for any market m is also estimated in this step. Let N^m denote the number of businesses purchasing advertising in market m, and let q_i^m, for $i = 1, 2, \ldots, N^m$, denote the quantity purchased by each of those businesses. We estimate the purchase density function in market m by using a kernel estimator,

$$\hat{g}^m(q) = \frac{1}{N^m h} \sum_{i=1}^{N^m} K(\frac{q - q_i^m}{h}) \tag{24}$$

for $q \in (0, \overline{q})$, where $K(\cdot)$ is a symmetric kernel function with compact support and h is a bandwidth. The purchase cumulative distribution function $\hat{G}^m(q)$ of market m can then be estimated by integrating $\hat{g}^m(q)$ in (24).

Following Lemma 1 , businesses' type $\theta^m(\cdot)$ can be estimated from $\theta^m(q) = 1/\hat{\xi}^m(q)$ and the marginal base intrinsic utility function $V_q(\cdot)$ is estimated from $\hat{V}_q(q) = \hat{T}^\dagger_q(q)\hat{\xi}^\dagger(q)$ by picking a benchmark market indexed by † under Assumption 3, where $\hat{\xi}^m(q)$ is estimated from,

$$\hat{\xi}^m(q) = \Big[1 - \hat{G}^m(q)\Big]^{1 - \frac{\hat{C}_Q(Q^m)}{\hat{T}^m_q(q)}} \exp\Big\{\hat{C}_Q(Q^m) \int_q^{\overline{q}} \frac{\hat{T}^{se}_{zz}(z)}{\hat{T}^m_z(z)^2} \log\Big[1 - \hat{G}^m(z)\Big] dz\Big\}, \tag{25}$$

where all terms on the right-hand side have been estimated or can be constructed from estimated terms. Once the marginal base intrinsic utility function $V_q(\cdot)$ is estimated, the base intrinsic utility function is then estimated from $\hat{V}(q) = \int_0^q \hat{V}_q(z)dz$ with the boundary condition $V(0) = 0$.

Third, we estimate the payoff function from market-level network effects for other markets, $W(\cdot, Q^m)$ for $m \neq \dagger$, by exploiting differences in equilibrium price schedules and purchases across markets. Recall that the payoffs from market-level network effects are normalized to zero. Following Proposition 4, we estimate $W(\cdot, Q^m)$ for any market m as follows:

$$\hat{W}(q, Q^m) = \int_0^q [\hat{T}_z^m(z)\hat{\xi}^m(z) - \hat{T}_z^\dagger(z)\hat{\xi}^\dagger(z)]dx, \tag{26}$$

where all terms on the right-hand side have been estimated or can be constructed from estimated terms.

Finally, we estimate the business taste type density functions $f^m(\cdot)$ and its aggregation function $\rho^m(\cdot,\cdot,\cdot,\cdot)$ for any market m using the estimated business taste type functions $\hat{\theta}^m(q)$. In particular, using the estimated business taste type functions $\hat{\theta}^m(q)$ and the purchases q_i^m, $i = 1, 2, \ldots, N^m$ in market m, we construct a pseudo taste type sample for the market by $\{\hat{\theta}_i^m = \hat{\theta}^m(q_i^m)\}_{i=1,2,\ldots,N^m}$. We then estimate $f^m(\cdot)$ using a kernel estimator,

$$\hat{f}^m(\theta) = \frac{1}{N^m h}\sum_{i=1}^{N^m} K(\frac{\theta - \hat{\theta}_i^m}{h}), \tag{27}$$

where $\theta \in (\underline{\theta}, \overline{\theta})$, $K(\cdot)$ is a symmetric kernel function with compact support and h is a bandwidth.

Moreover, we can use the estimated pseudo sample of the business taste type $\{\hat{\theta}_i^m\}_{i=1,2,\ldots,N^m}$, type mapping $\hat{\theta}^m(q)$, and observables $\{x_i^m, x_j^m\}_{i=1,2,\cdots,N^m; j=1,2,\cdots,J}$ from multiple markets to directly estimate the aggregation function $\rho(\cdot,\cdot,\cdot,\cdot)$ following our identification strategy.

Discussion of Estimator Properties

Our first step is a simple regression to construct the tariff functions and the quality-adjusted quantity. Because we assume the analyst knows the price schedule, we treat the tariff functions and quantities as data. We now briefly discuss the properties of the estimators in the latter steps. The marginal cost for market m exhibits a strong consistency following Galambos (1978) and the standard delta method. Following the empirical process literature (see, e.g., Andrews (1994)), we can view $\widehat{V}_q(\cdot)$ and $\theta^m(\cdot)$ as stochastic processes defined on

$[\underline{q}, \overline{q}]$. Because they are smooth functionals of the empirical c.d.f., one can establish their uniform almost sure convergence at a parametric rate. See Luo, Perrigne and Vuong (2018) for a detailed discussion. Similar arguments apply to $\widehat{W}(\cdot, Q^m)$. Finally, because the pseudo types converge at a faster rate than the optimal rate for kernel density estimation, our kernel estimator has the standard asymptotic properties of uniform convergence.

5 Empirical Application

5.1 Data on the Yellow Pages Print Industry

In our empirical application, we apply our estimation approach to yellow pages print advertisement industry in 2011 for investigating the importance of both industry-level and market-level network effects. Although the usage of print yellow pages directories has been declining as both businesses and consumers increasingly turning to internet search engines and online directories, there are four main reasons for which we choose to examine this industry.

First, the yellow pages directory is one of traditional media platforms exhibiting network effects. Businesses and consumers interact with each other via yellow pages directories. Consumers value yellow pages directories for information, and businesses value yellow pages directories as a way to advertise to consumers. More advertising leads to more consumer usage, which in turn leads to more advertising such that the interaction between consumers and businesses create a network effects. Various other important media platforms share features similar to that of the yellow pages directories, including other print media such as newspapers and magazines, online yellow pages directories, internet search advertising platforms such as Google Ads, Bing/Yahoo, Amazon. Our approach sheds light on examining the importance of network effects for these platforms.

Second, the yellow pages print industry was still a prominent advertising industry by 2011. According to a study by the Yellow Pages Association, 75 percent of adults in the United States still used print yellow pages, and for every \$1 in investment, businesses returned \$15, in February 2011. BIA/Kelsey estimated that the revenues from print yellow pages directories globally. The online yellow pages directories also became a popular media platform in 2011. According to several reports, the search term "yellow pages" was in the top 5 highest revenue generators of all search terms in Google's AdWords program in 2010. Experian/Hitwise reported "yellow pages" as being one of the most searched for terms on the Internet in 2011.

Third, we examine a model relying on an important assumption which requires a monopoly platform. In practice, it is rare to have only one platform in a two-sided market. For example, there are several Internet search advertising platforms, and the yellow pages directories in a district of the U.S. are usually published by two or three different publishers. In our application, we examine the yellow page directories of Toronto, published by a monopoly firm, which satisfy our model assumption.

Fourth, our empirical approach requires micro-level data on purchase quantities and their prices from multiple markets. We hand-collect advertising quantity (size, color, and other features) purchased by each business and obtain the corresponding price schedules from the Yellow Pages Association. In comparison, the purchase quantity and price data for prevailing platforms, such as newspapers and magazines, online yellow pages directories, and Internet search advertising platforms, are often not available.

We collect data on 2011 print advertisements in the yellow pages directories for the city of Toronto.[20] In Toronto, there is only one publisher of the print yellow pages directory, the Yellow Pages Group. The company publishes print yellow pages directories for each of the seven non-overlapping districts of Toronto and distributes them freely to the households living in the corresponding districts. The map in Figure (1) shows these seven districts, which are Core Center (C), Core West (W), Core Northeast (NE), Core Southeast (SE), Etobicoke (E), North (N) and Scarborough (S). We treat each of the seven districts as separate markets in our empirical application.

The publisher collects revenue by selling advertisement slots in these yellow pages directories to local businesses. The price businesses pay depends on the size and categories they use. We collect our data through two sources: (i) we manually read off businesses' advertisement purchases from the directories; (ii) the price schedule data for each directory is collected from the Local Search Association.[21] Advertisements in yellow pages directories differ by size, color and other special features, which provide us with numerous possible combinations of advertising options for businesses to choose from. For instance, we observe 212 and 225 advertising options chosen by businesses, which lead to corresponding numbers of different prices paid by businesses in SE and NE yellow pages directories, respectively. Following Busse and Rysman (2005), we estimate a nonlinear tariff function for each directory and

[20] An increasing competitor to print yellow pages advertising is internet search engines and internet yellow pages advertising. The print yellow pages is predicted to go extinct in the future. However, the industry still remains very strong. As of 2007, 71% of Canadians still consulted print yellow pages directories every month (2007 Canadian Business Usage Study), and 30 million print directories were distributed annually in Canada (comScore Media Metrix (Nov. 2007)). Moreover, according to the Newsletter by Local Search Association (2011), the print yellow pages remain the people's first choice to search for local businesses.

[21] The name of the association was "Yellow Pages Association" prior to 2011.

construct quality-adjusted quantities for advertisement purchases.

Regarding price schedules, there are two interesting features worth noting. The first is that the schedules display a nonlinear pattern, as reported in earlier papers (e.g., Busse and Rysman (2005) and Aryal and Gabrielli (2020)). The more advertisement space businesses buy, the more discount they get. For instance, for multi-color advertisements in the NE yellow pages directory, the price per square pica varies from $11.1 to $9.22 and $4.48 for the smallest to half-page and double-page advertisements, respectively. The same pattern is also observed in other categories. This feature motivates us to employ a nonlinear pricing model.

The second feature is that prices and discounts vary across markets, which will provide us with important variation to identify our model primitives as we will discuss later. For example, the prices are $1935.60, $13697.40 and $27049.20 for the smallest, half-page and double-page multi-color advertisements in the SE yellow pages directory, while corresponding prices are $2304.00, $16381.80 and $32371.20 in the NE yellow pages directory. It is therefore more expensive to advertise in the NE yellow pages directory than the SE one. On the other hand, there are more residents living in the NE district than SE. Moreover, as we will show later, a higher quantity of advertisements are purchased in the NE directory than that in the SE directory. These two facts suggest that there may exist network effects in the yellow pages advertising markets of Toronto, and the publisher exploits the network effects to charge a higher price for advertising in the NE directory.

As it is quite troublesome to manually collect businesses' advertisement space purchases from the directories, we read off purchase data from two directories, NE and SE, to illustrate our empirical methodology. We collect a total of 6904 advertisements over $1,458$ industry headings and $7,206$ advertisements over $1,455$ industry headings from SE and NE yellow pages directories, respectively. The publisher collected revenues of $5.38 million and $6.98 million from selling advertisement space in SE and NE directories, respectively. Tables (1) and (2) display the top 10 industry headings, which represent about 20% of the total revenue from both directories. These industries include professional and household services. Within each of these industry headings, businesses in industries with larger total size of advertisements tend to buy larger advertisements themselves. This feature suggests that there might exist interactions between businesses within industries, which motives us to analyze industry-level network effects.

5.2 Empirical Results

We apply our estimation approach to investigate the importance of both industry-level and market-level network effects by using advertising data on directories of SE and NE districts in Toronto. We first estimate the model primitives following our estimation approach and then use them to assess the businesses' information rents and payoffs from market-level network effects. Finally, we conduct counterfactual experiments to investigate the importance of both industry-level and market-level network effects.

5.2.1 Estimated Model Primitives

In view of our estimation approach, we need to estimate the tariff function and construct purchase distributions in the first step. We then pick the SE district as the benchmark market to estimate the base intrinsic payoff function and use advertising data from the NE district to estimate the payoff function from market-level network effects.[22] Finally, we estimate the business taste type density functions and its aggregation function using constructed pseudo samples of business types for both districts.

Using the price schedule data for multicolored display advertisements from all seven districts for yellow pages directories in Toronto, we estimate (21). The coefficient estimates are $\hat{\beta}_0 = -18.3742$ (2.0169), $\hat{\beta}_1 = 0.7547$ (0.0163), and $\hat{\beta}_2 = 1.7066$ (0.1530), where standard errors are in parentheses. Since all the estimates are significant and the adjusted R^2 of the regression is 0.9764, we use those regression estimates to construct quality-adjusted quantities and the tariff functions from (22) and (23). The first two rows of Tables (3) and (4) show the summary statistics of quality-adjusted quantities and tariffs paid for both SE and NE yellow pages directories, respectively. The publisher charges a higher unit price for advertising in the NE directory than in the SE directory.

The total quality-adjusted quantities purchased are 493.47 thousand and 448.15 thousand square picas in the NE and SE yellow pages directories, respectively. Using (23), we obtain an estimate for the marginal cost function $\hat{C}_Q(Q) = \hat{T}_q(\overline{q}, Q) = 9.9738 \times 10^{-10} \times Q^{1.7066}$, where $\overline{q}$ is the quality-adjusted maximum size of advertisement in our context. The marginal cost function is increasing with respect to the total quantity produced, which is consistent with our model Assumption 1-(iv). This may be because the publisher incurs more cost in designing the layout of a yellow pages directory to accommodate more advertisements.

We now report the estimates of the base intrinsic payoff function $V(\cdot)$ and the payoff

[22] We pick the SE district as the benchmark market for a convenient interpretation because the estimated network utility becomes positive.

function from market-level network effects for purchases in the NE yellow page directory, $W^{ne}(\cdot) \equiv W(\cdot, Q^{ne})$.[23] These functions are displayed in Figure 2. The estimated base intrinsic payoff function is increasing and concave, and the estimated payoff function from market-level network effects is positive, increasing, and concave. These estimated payoff functions ensure that Assumption 1 -(i) is satisfied. As we can observe from the graph, the payoff function from market-level network effects is quite relevant, as its scale is comparable with a median consumer's utility function. The positive market-level network effects we find is consistent with Rysman (2004). A larger yellow pages directory contains more information that is more valuable for consumers, thereby leading to more usage, which in turn increases the value of advertising for businesses.

We finally report the pseudo sample estimates of the business taste types in the SE and NE districts, which are used to estimate its probability density functions and aggregation function capturing industry-level network effects. Summary statistics on estimated pseudo business types are displayed in the middle of Tables 3 and 4. Figure 3 displays the estimated density functions $f^{se}(\cdot)$ and $f^{ne}(\cdot)$. Businesses in the NE district tend to have higher tastes for yellow pages advertisements. We interpret this result by positing that there may be more residential consumers or residents who have higher income in the NE district such that businesses expect advertising is the NE directory to be more valuable.

Recall that business taste type θ aggregates all information on which the publisher cannot price discriminate, including a business' individual-level and industry-level characteristics, and the mean advertisement purchase quantity in the business' industry; i.e., $\theta = \rho(x_i, x_j, \bar{q}_j) + \nu_i$ for a business i in industry j. It captures the industry-level network effects through the mean purchase quantity $\bar{q}_j$. In principle, we can estimate the function nonparametrically from the estimated pseudo samples of θ and observed individual-level and industry-level characteristics, and the mean purchase quantity in each industry using standard nonparametric approaches. However, the nonparametric approach often requires a large sample on observables. See Li and Racine (2007). Due to data limitations, we adopt a parametric approach for this step. Since $\theta \in [0, 1]$ by our normalization in Assumption 3, we first transform θ into an unbounded variable $\vartheta = \log \frac{\theta}{1-\theta}$, which is a strictly increasing one-to-one mapping. It implies that $\theta = e^{\vartheta}/(1 + e^{\vartheta})$. We then estimate the following regression equation,

$$\vartheta_i = \alpha_0 + \text{Competition}_i' \alpha_1 + X_i' \alpha_2 + \nu_i$$

where Competition_i is a vector of variables regarding industry-level competition, including

[23] Note that we have picked SE as the benchmark market such that the payoffs from market-level network effects for businesses purchasing SE advertising are all zero.

industry-level fixed effect and mean purchase quantity, X_i is a vector of individual characteristics, and ν_i is the unobservables with mean zero conditional on Competition and observables X. The mean zero assumption directly follows our identification and is commonly made in the literature on social interaction with the reflection problem.[24]

In this study, since we do not observe businesses' individual-level characteristics, we include only the number of firms and industry-level fixed effects for industry-level characteristics in the regression.[25] Table 5 displays the results of our regression with four specifications. Note that the adjusted R^2 are very similar, while the squared terms and the number of firms do not significantly affect ϑ. Thus, Specification I is preferred as the other variables are not significant and do not add much explanatory power. The results show that the more advertising a business' competitors purchase from the yellow pages directory, the higher the business' taste tends to be. Therefore, there are positive network effects at the industry level. To make its advertisement stand out, a business tends to buy a larger advertisement than that of its competitors.

5.2.2 Evaluating the Payoffs of the Publisher and Businesses

Using the above estimated primitive estimates, we empirically evaluate the publisher's revenue as well as the businesses' informational rents and payoffs from market-level network effects. The publisher's revenue is the total payment she collects from businesses. The publisher collects 5.38 million and 6.98 million dollars from selling advertising in SE and NE directories, respectively. The informational rent is defined as the difference between utility and payment. For instance, consider a type-θ business who purchases q quantity of advertisement in the NE directory by paying $T^{ne}(q)$ and obtains utility $U(q, \theta, Q^{ne})$. The informational rent for this business is given by $U(q, \theta, Q^{ne}) - T^{ne}(q)$. The total informational rents are 4.82 million and 6.26 million dollars for businesses buying advertising space in the SE and NE directories, respectively. The ratio of total informational rents to total payment across all consumers measures the cost of asymmetric information, which is 89.26% in SE and 90.20% in the NE. Although the publisher collects more revenue, businesses also obtain a substantial amount of informational rent due to their private information. Compared to the results reported in Aryal and Gabrielli (2020), businesses in Toronto obtain more informational rent. This comes from the fact that businesses display a higher degree of heterogeneity in Toronto than State College, Pennsylvania. The publisher leaves more

[24] See e.g., Manski (1993).

[25] There are more than a thousand different industry headings. As it is unrealistic to have this many fixed effects in our regression, we thus define 10 division dummies using SIC code, such as agriculture, forestry, and fishing, mining, construction, manufacturing, and so on.

informational rent to screen businesses of different tastes. Tables (3) and (4) show the summary statistics of taste, informational rent, and its ratio to payment.

Regarding the payoffs from market-level network effects, businesses purchasing advertisements in the NE directory obtain a total payoff from market-level network effects that is 2.96 millions dollars higher than those purchasing advertising in the SE directory. This amount accounts for about 22.36% of the total payoffs. It is also about 42.41% of total payment and 47.28% of the total informational rent that all businesses purchasing advertising space in the NE directory obtain. Both the publisher and businesses benefit from market-level network effects. Table (4) displays the summary statistics of the network payoff obtained by each business.

5.2.3 Counterfactuals

With estimated model primitives at hand, we conduct counterfactuals to evaluate the welfare implications of the nonlinear pricing scheme. Nonlinear pricing is prevalent in many industries, such as advertising, telecommunications, electricity, food, and health care. While it increases the profits of sellers with market power, its welfare effects are distributional and often subject to policy debate. In particular, nonlinear pricing disproportionately impacts low-type buyers through exclusion and higher unit prices. Because lower-income people tend to be low types, nonlinear pricing may limit their access to essential goods and services. See, e.g., Attanasio and Frayne (2006) for food pricing in Colombia, Borenstein (2012) for electricity pricing in California, Dalton (2014) for health insurance pricing, and Luo, Perrigne and Vuong (2018) for mobile service pricing. Our counterfactuals compare prices, quantities, business surplus, and publisher profit under two pricing schemes with and without network effects. As far as we know, our paper is the first to empirically quantify the distributional impacts of nonlinear pricing when network effects play important roles.

Our counterfactual exercises consist of two parts. First, we conduct counterfactuals to evaluate the importance of network effects under the nonlinear pricing scheme by shutting down the market-level and/or industry-level network effects. Second, we evaluate the welfare implications of the nonlinear pricing scheme by simulating the equilibrium outcomes under an alternative linear pricing scheme. Under the linear pricing scheme, price discrimination is not allowed and the publisher can only charge one flat price to all businesses. We evaluate the welfare implications of the nonlinear pricing scheme by comparing its resulting purchase quantities, publisher revenue, and business surpluses with those under the linear pricing scheme. As we normalize the payoffs from market-level network effects to zero for all businesses that purchase advertising in the SE directory, our counterfactual experiments focus on the NE directory.

To conduct counterfactual experiments, we need to make some assumptions on the elements we do not estimate. As we identify and estimate the market-level network effects $W(\cdot, Q)$ only for $Q = Q^{ne}$, we assume that for any $Q \in [\underline{Q}, \overline{Q}]$,

$$\hat{U}(q, \theta, Q) = \theta\Big[\hat{V}(q) + (Q - Q^{se}) \times \hat{W}_{ne}(q)\Big], \tag{28}$$

which equals the estimated intrinsic payoff when $Q = Q^{se}$, and equals the sum of the estimated intrinsic payoff and the payoff from market-level network effects for the NE directory when $Q = Q^{ne}$. Regarding the cost function, we assume the fixed cost to be zero. Note that we identify only the marginal cost $C_Q(\cdot)$. However, if Q does not change a lot, the publisher's optimal decision only involves the marginal cost $C_Q(\cdot)$. Assuming zero fixed cost would not change the result on business surplus or publisher revenue.

NONLINEAR PRICING WITH ONLY INDUSTRY-LEVEL NETWORK EFFECTS

Our first counterfactual experiment simulates equilibrium outcomes in a model under the nonlinear pricing with only industry-level network effects. We assume businesses do not enjoy market-level network effects in this case. Formally, we rewrite their utility as

$$U^{new}(q, \theta, Q) = \theta \hat{V}(q).$$

We then use $(\hat{C}(\cdot), U^{new}(\cdot,\cdot,\cdot), \hat{f}^{ne}(\cdot))$ to simulate the equilibrium market outcomes in this case, which is defined by

$$\begin{aligned} \theta \hat{V}_q(q) &= \hat{C}_Q(Q) + \frac{1 - \hat{F}^{ne}(\theta)}{\hat{f}^{ne}(\theta)} \hat{V}_q(q), \\ T_q(q) &= \theta \hat{V}_q(q), \end{aligned}$$

along with two boundary conditions: (i) $q(\underline{\theta}) = 0$; (ii) $T(0) = 0$ as well as a self-fulfilling expectation condition: $\int_{\underline{\theta}}^{\overline{\theta}} q(\theta) d\theta = Q$. We remark that $q(\cdot)$ is defined as the inverse mapping of $\theta(\cdot) \equiv \hat{H}^{-1}[\hat{C}_Q(Q)/\hat{V}_q(\cdot)]$, where $\hat{H}(\cdot) \equiv \cdot - \frac{1-\hat{F}(\cdot)}{\hat{f}(\cdot)}$ is increasing and $\hat{V}_q(\cdot)$ is decreasing. Thus, the implied $\theta(\cdot)$ is increasing. To find the new equilibrium total purchase quantity Q, we use the bisection method. In particular, for any value of Q, we can solve the above two equations and find the optimal assignment schedule $q(\cdot)$ and tariff function $T(\cdot)$. We then use the bisection method to search for the Q such that $\int_{\underline{\theta}}^{\overline{\theta}} q(\theta) \hat{f}^{ne}(\theta) d\theta = Q$. As there are only industry-level network effects in this case, we denote the equilibrium functions for purchases and tariffs by $q_n^{ind}(\cdot)$ and $T_n^{ind}(\cdot)$, respectively. In addition, the equilibrium purchases and tariff functions observed in the data with both industry-level and market-level network effects are denoted by $q_n^{full}(\cdot)$ and $T_n^{full}(\cdot)$, respectively.

Nonlinear pricing with no network effects

Our second counterfactual experiment simulates the model equilibrium outcomes under the nonlinear pricing scheme with no network effects. We shut down both the industry-level and market-level network effects in this model. The market-level network effects are shut down using the same method in the first experiment. The industry-level network effects are reflected by the effects of industry-level fixed effects and purchases. To shut down industry-level network effects, we calculate a new ϑ_i as $\vartheta_i^0 = \vartheta_i - \text{Competition}_i'\hat{\alpha}_1$. We then calculate a new taste using $\theta_i^0 = \exp(\vartheta_i^0)/(1 + \exp(\vartheta_i^0))$, which represents the business' taste for advertising in the yellow pages without pressure from its competitors. We estimate a new density function $f^{new}(\cdot)$ using the new type sample $\{\theta_i^0\}_{i=1,2,\ldots,N}$. Table (4) gives summary statistics on θ^0, and Figure (3) gives a kernel estimate of the new density function for businesses that purchase advertising in the NE directory. After shutting down industry-level network effects, firms tend to have less willingness-to-pay for advertising in the yellow pages directory. Thus, we strengthen the findings in Rysman (2004) by finding positive network effects at the industry level.

To calculate the effects of shutting down all network effects, we simulate the market outcome using $(\hat{C}(\cdot), U^{new}(\cdot,\cdot,\cdot), f^{new}(\cdot))$. As there are no network effects and the pricing schedule is nonlinear in this model, we denote the optimal purchases and tariff functions by $q_n^0(\cdot)$ and $T_n^0(\cdot)$, respectively.

Linear pricing with both industry-level and market-level network effects

Our third counterfactual experiment simulates equilibrium outcomes under the linear pricing scheme in a full model with both industry-level and market-level network effects. Recall that price discrimination is not allowed and the publisher can only set one flat price in this case. To ensure that the utility of a type-θ business is the same as the utility under the nonlinear pricing scheme given by (28), we restrict the publisher's total production to the observed Q^{ne} in the NE directory. The publisher chooses a price to maximize her profit and all businesses choose a purchase quantity to maximize their payoff. Moreover, the total purchased quantity is restricted to Q^{ne} in equilibrium.

In particular, the new equilibrium in this case is defined by

$$\max_{p} \int_{\underline{\theta}}^{\overline{\theta}} pq(\theta)\hat{f}^{ne}(\theta)d\theta - \hat{C}(Q^{ne}),$$
$$\text{s.t.} \quad \hat{U}_q(q,\theta,Q^{ne}) = p, \quad \forall\, \theta \in [\underline{\theta},\overline{\theta}],$$

and $\int_{\underline{\theta}}^{\overline{\theta}} q(\theta)\hat{f}^{ne}(\theta)d\theta = Q^{ne}$, where p is the price for one unit of advertisement, and estimated utility function $\hat{U}(q,\theta,Q^{ne})$ is given by (28). Note that $q(\cdot) = \hat{U}_q^{-1}(p,\cdot,Q^{ne})$ for any p in equilibrium. To find the equilibrium, we solve for the optimal price p and purchase quantity $q(\cdot)$ for Q^{ne} such that $\int_{\underline{\theta}}^{\overline{\theta}} q(\theta)\hat{f}^{ne}(\theta)d\theta = Q^{ne}$. Given that there are both industry-level and market-level network effects and that the pricing schedule is linear in this model, we denote the equilibrium purchases and tariffs in this case by $q_l^{full}(\cdot)$ and $T_l^{full}(\cdot)$, respectively.

Linear pricing with only industry-level network effects

Our fourth counterfactual experiment simulates equilibrium outcomes in a model under linear pricing with only industry-level network effects. In this case, businesses do not enjoy the market-level network effects such that the utility for a type-θ business is given by $\theta\hat{V}(q)$. The new equilibrium in this case is then defined by

$$\max_{p} \int_{\underline{\theta}}^{\overline{\theta}} pq(\theta)\hat{f}^{ne}(\theta)d\theta - \hat{C}(Q),$$
$$\text{s.t.} \quad \theta\hat{V}_q(q) = p, \quad \forall\, \theta \in [\underline{\theta},\overline{\theta}].$$

The equilibrium $q(\cdot)$ is defined by $q(\cdot) = \hat{V}_q^{-1}(p/\cdot)$ in this case. To find the equilibrium, we first solve for the optimal price p and purchase quantity $q(\cdot)$ for any value of Q. We then search for the equilibrium Q such that $\int_{\underline{\theta}}^{\overline{\theta}} q(\theta)\hat{f}^{ne}(\theta)d\theta = Q$. As there are only industry-level network effects and the pricing schedule is linear in this model, we denote the equilibrium purchases and tariff functions in this case by $q_l^{ind}(\cdot)$ and $T_l^{ind}(\cdot)$, respectively.

Linear pricing with no network effects

Our final counterfactual experiment simulates equilibrium outcomes in a model under linear pricing with no network effects. We shut down both the industry-level and market-level network effects in this case. The market-level network effects are shut down using the same method as in the first experiment. The industry-level network effects are shut down using the new taste distribution $f^{new}(\theta^0)$ as in the second experiment. The new equilibrium in this case is then defined by

$$\max_{p} \int_{\underline{\theta}^0}^{\overline{\theta}^0} pq(\theta^0)\hat{f}^{new}(\theta^0)d\theta^0 - \hat{C}(Q),$$
$$\text{s.t.} \quad \theta^0\hat{V}_q(q) = p, \quad \forall\, \theta^0 \in [\underline{\theta}^0, \overline{\theta}^0].$$

The equilibrium $q(\cdot)$ is defined by $q(\cdot) = \hat{V}_q^{-1}(p/\cdot)$ in this case. To find the equilibrium, we first solve for the optimal price p and purchase quantity $q(\cdot)$ for any value of Q. We then search for the equilibrium Q such that $\int_{\underline{\theta}^0}^{\overline{\theta}^0} q(\theta^0)\hat{f}^{new}(\theta^0)d\theta^0 = Q$. As there are no network effects and the pricing schedule is linear in this model, we denote the equilibrium purchases and tariff function in this case by $q_l^0(\cdot)$ and $T_l^0(\cdot)$, respectively.

Discussions on Counterfactual Results

Our counterfactuals assess the overall and distributional welfare effects of network effects associated with both nonlinear and linear pricing schemes. Nonlinear pricing generally hurts low-type consumers the most through exclusion and higher unit prices. See, e.g., Luo, Perrigne and Vuong (2018). Similar to Aryal and Gabrielli (2020), a substantial amount of businesses purchase no advertisement in our sample; these businesses are excluded under both schemes. Therefore, the yellow page industry is less ideal in demonstrating the exclusion effects of nonlinear pricing. Instead, we assume the publisher covers the same range of businesses purchasing advertisements under nonlinear pricing and investigate how different types in this range are affected in alternative settings. We measure the business surplus, publisher profit, total welfare (as the sum of the two), payment, and purchase quantity with different network effects for each pricing scheme. We then quantify the distributional impacts of nonlinear pricing on businesses of different types.

Table (6) summarizes our counterfactual results. Table (7) further assesses the distributional impacts of nonlinear pricing by dividing the sample into three equal subsamples from the lowest to the highest types of businesses. The linear pricing columns report the actual values, while the nonlinear pricing columns correspond to nonlinear pricing as proportions of the linear pricing values.

Table (6) shows that nonlinear pricing generally makes businesses worse off but makes the publisher better off as the total business surplus decreases by 8.4% and the publisher's profit increases by 17.2% in the model with both market- and industry-level network effects. This is due to businesses paying more under nonlinear pricing, though they tend to consume more. Moreover, both businesses and the publisher benefit significantly from network effects.

The network effects at both market and industry levels increase the total business surplus by 60.62% (from 4.34 million to 6.98 million) and increase the total publisher profit by 101.50% (from 2.50 million to 5.03 million), which leads to an increase of 75.53% in total welfare (from 6.8 million to 12.01 million).

Table (7) shows that the low-type group is most hurt by nonlinear pricing because of higher unit prices. Higher unit prices lead businesses in this group to purchase fewer products and receive fewer surpluses than those under linear pricing. Take the model with both market- and industry-level network effects as an example. The unit price for the low-type group is over three times what it is under linear pricing, while the corresponding value is 116% for the high-type group. As a result, the purchase quantity and surplus per business in the low-type group are only 27% and 40% of those under linear pricing, and the corresponding values in the high-type group are 133% and 97%. We also find that the network effects play important roles in shaping the distributional effects of nonlinear pricing. Without network effects, nonlinear pricing hurts low-type businesse more. The purchase quantity and surplus per business in the low-type group are only 9% and 13% of what they are under linear pricing.

6 Conclusion

This paper examines the importance of nonlinear pricing and network effects in determining the welfare of market participants. We first propose a novel approach to identify and estimate a nonlinear pricing model featuring both industry-level and market-level network effects by exploiting the FOCs of both sides of the market. Our empirical approach is applied to investigate the yellow pages print industry. With hand-collected data on advertisers' purchases and nonlinear price schedules from seven districts in Toronto, we estimate a model that allows for heterogeneous business demand and endogenizes the publisher's optimal nonlinear pricing decision. We find positive network effects at both the industry and market levels. Our counterfactuals evaluate the overall and distributional welfare effects of the nonlinear pricing relative to linear pricing with and without network effects.

The methodology we develop in this paper is also applicable to other industries featuring network effects, such as broadband internet, cellular phone service, and social media advertising. It is also useful to study technology adoption with network externalities. While the divide-and-conquer strategy simplifies our setting, incorporating interactions between consumers, businesses, and the platform may become necessary in other settings. Endogenizing all these three players' decisions is challenging, yet potentially useful to study platforms, such as online dating websites, online advertising platforms.

References

Andrews, D.W., 1994. Empirical process methods in econometrics, Elsevier. volume 4 of *Handbook of Econometrics*, pp. 2247–2294.

Armstrong, M., 1996. Multiproduct nonlinear pricing. Econometrica 64, 51–75.

Armstrong, M., Vickers, J., 2001. Competitive price discrimination. The RAND Journal of Economics 32, 579–605.

Aryal, G., Gabrielli, M.F., 2020. An empirical analysis of competitive nonlinear pricing. International Journal of Industrial Organization 68, 102538.

Ascarza, E., Lambrecht, A., Vilcassim, N., 2012. When talk is “free”: The effect of tariff structure on usage under two-and three-part tariffs. Journal of Marketing Research 49, 882–899.

Attanasio, O., Frayne, C., 2006. Do the poor pay more? Working Paper .

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

Blume, L.E., Brock, W.A., Durlauf, S.N., Ioannides, Y.M., 2011. Identification of social interactions, in: Handbook of social economics. Elsevier. volume 1, pp. 853–964.

Blume, L.E., Brock, W.A., Durlauf, S.N., Jayaraman, R., 2015. Linear social interactions models. Journal of Political Economy 123, 444–496.

Borenstein, S., 2012. The redistributional impact of nonlinear electricity pricing. American Economic Journal: Economic Policy 4, 56–90.

Bousquet, A., Ivaldi, M., 1997. Optimal pricing of telephone usage: An econometric implementation. Information Economics and Policy 9, 219–239.

Bramoullé, Y., Djebbari, H., Fortin, B., 2009. Identification of peer effects through social networks. Journal of Econometrics 150, 41–55.

Brock, W.A., Durlauf, S.N., 2001a. Discrete choice with social interactions. The Review of Economic Studies 68, 235–260.

Brock, W.A., Durlauf, S.N., 2001b. Interaction-based models, Elsevier, Amsterdam, pp. 3297–3380.

Brock, W.A., Durlauf, S.N., 2007. Identification of binary choice models with social interactions. Journal of Econometrics 140, 52–75.

Busse, M., Rysman, M., 2005. Competition and price discrimination in yellow pages advertising. The Rand Journal of Economics 36, 378–390.

Chen, L., Luo, Y., Xiao, P., 2019. Estimating the demand for service bundles under three-part tariffs. The Journal of Industrial Economics 67, 448–483.

Dalton, C.M., 2014. Estimating demand elasticities using nonlinear pricing. International Journal of Industrial Organization 37, 178–191.

Durlauf, S.N., Ioannides, Y.M., 2010. Social interactions. Annu. Rev. Econ. 2, 451–478.

Economides, N., 1996. Network externalities, complementarities, and invitations to enter. European Journal of Political Economy 12, 211–233.

Ekeland, I., Heckman, J.J., Nesheim, L., 2004. Identification and estimation of hedonic models. Journal of Political Economy 112, 60.

Galambos, J., 1978. The Asymptotic Theory of Extreme Order Statistics. New York: John Wiley.

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

Grubb, M.D., 2009. Selling to overconfident consumers. American Economic Review 99, 1770–1807.

Halaburda, H., Yehezkel, Y., 2011. Platform competition under asymmetric information. NET Institute Working Papers .

Heckman, J.J., Matzkin, R.L., Nesheim, L., 2010. Nonparametric identification and estimation of nonadditive hedonic models. Econometrica 78, 1569–1591.

Huang, C.I., 2008. Estimating demand for cellular phone service under nonlinear pricing. Quantitative Marketing and Economics 6, 371.

Lambrecht, A., Seim, K., Skiera, B., 2007. Does uncertainty matter? consumer behavior under three-part tariffs. Marketing Science 26, 698–710.

Lee, L.F., 2007. Identification and estimation of econometric models with group interactions, contextual factors and fixed effects. Journal of Econometrics 140, 333–374.

Li, Q., Racine, J.S., 2007. Nonparametric econometrics: theory and practice. Princeton University Press.

Luo, Y., 2023. Bundling and nonlinear pricing in telecommunications. The RAND Journal of Economics 54, 268–298.

Luo, Y., Perrigne, I., Vuong, Q., 2018. Structural analysis of nonlinear pricing. Journal of Political Economy 126, 2523–2568.

Manski, C., 1993. Identification of endogenous social effects: The reflection problem. The Review of Economic Studies 60, 531–542.

Manski, C.F., 2000. Economic analysis of social interactions. Journal of Economic Perspectives 14, 115–136.

Maskin, E., Riley, J., 1984. Monopoly with incomplete information. The Rand Journal of Economics 15, 171–196.

Menzel, K., 2016. Inference for games with many players. The Review of Economic Studies 83, 306–337.

Miravete, E.J., Röller, L.H., 2004. Estimating price-cost markups under nonlinear pricing competition. Journal of the European Economic Association 2, 526–535.

Moffitt, R.A., 2001. Policy interventions, low-level equilibria, and social interactions. Social Dynamics 4, 6–17.

Nevo, A., Turner, J.L., Williams, J.W., 2016. Usage-based pricing and demand for residential broadband. Econometrica 84, 411–443.

Rochet, J.C., Stole, L.A., 2002. Nonlinear pricing with random participation. The Review of Economic Studies 69, 277–311.

Rysman, M., 2004. Competition between networks: A study of the market for yellow pages. Review of Economic Studies 71, 483–512.

Sweeting, A., 2009. The strategic timing incentives of commercial radio stations: An empirical analysis using multiple equilibria. The RAND Journal of Economics 40, 710–742.

Xu, H., 2018. Social interactions in large networks: A game theoretic approach. International Economic Review 59, 257–284.

Appendix

1. Proofs

Proof of Proposition 1

Proposition 1.

(i) *Under Assumption 1, the publisher's optimal fulfilled expectation price contract,* $[q^*(\cdot), T^*(\cdot)]$ *satisfies the necessary conditions,*

$$U_q(q^*(\theta), \theta, Q^*) = C'(Q^*) + \frac{1 - F(\theta|\bar{\boldsymbol{q}}^*)}{f(\theta|\bar{\boldsymbol{q}}^*)} U_{q\theta}(q^*(\theta), \theta, Q^*), \tag{29}$$

$$T_q(q^*(\theta)) = U_q(q^*(\theta), \theta, Q^*) \tag{30}$$

for all $\theta \in [\underline{\theta}, \overline{\theta}]$*, where* $\bar{\boldsymbol{q}}^* = (\bar{q}_1^*, \cdots, \bar{q}_J^*)$ *and* Q^* *are the mean industry-level purchases and total market-level purchase fulfilled in equilibrium.*

(ii) *If* $U_q(q, \theta, Q)$ *is bounded for all* q, θ, Q*, then the optimal fulfilled expectation price schedule always exists.*

Proof. **Part** (i)**:** Part (i) consists of two steps. First, we show the necessary conditions (29) and (30) that the publisher's optimal price schedule must satisfy for any expectations $\bar{\boldsymbol{q}}^e$ and Q^e. Second, we enforce the fulfilled conditions that the expectations $\bar{\boldsymbol{q}}^e$ and Q^e should satisfy in equilibrium.

Step 1**:** Given any expectations $\bar{\boldsymbol{q}}^e$ and Q^e, the publisher's optimal price contract solves the following constrained (expected) profit maximization problem:

$$\max_{q(\cdot), T(\cdot)} \int_{\underline{\theta}}^{\overline{\theta}} T(q(\theta)) f(\theta|\bar{\boldsymbol{q}}^e) d\theta - C(Q^e), \tag{31}$$

$$\text{s.t. (IC)} \quad \theta = \underset{x}{\text{argmax}} \quad U(q(x), \theta, Q^e) - T(q(x)), \quad \forall\, \theta \in [\underline{\theta}, \overline{\theta}]. \tag{32}$$

Note that we assume full coverage of businesses such that businesses' individual rationality condition for the lowest type is always satisfied. The derivation in this step has four parts.

(a) Reduction of IC (32) **to a simpler form.** The necessary and sufficient conditions for IC (32) are given by

$$U_q(q(\theta),\theta,Q^e)q_\theta(\theta) - T_q(q(\theta))q_\theta(\theta) = 0, \tag{33}$$

$$U_{qq}(q(\theta),\theta,Q^e)q_\theta^2(\theta) + U_q(q(\theta),\theta,Q^e)q_{\theta\theta}(\theta) - T_{qq}(q(\theta))q_\theta^2(\theta) - T_q(q(\theta))q_{\theta\theta}(\theta) \le 0. \tag{34}$$

Differentiating (33) with respect to θ yields

$$T_{qq}(q(\theta))q_\theta^2(\theta) + T_q(q(\theta))q_{\theta\theta}(\theta) = U_{q\theta}(q(\theta),\theta,Q^e)q_\theta(\theta) + U_{qq}(q(\theta),\theta,Q^e)q_\theta^2(\theta) + U_q(q(\theta),\theta,Q^e)q_{\theta\theta}(\theta),$$

which we substitute into (34) to give

$$U_{q\theta}(q(\theta),\theta,Q^e)q_\theta(\theta) \ge 0.$$

The assumption ensures that $U_{q\theta}(q(\theta),\theta,Q^e) > 0$, which implies that the IC condition (32) can be reduced to:

$$U_q(q(\theta),\theta,Q^e) = T_q(q(\theta)) \tag{35}$$

$$q_\theta(\theta) \ge 0. \tag{36}$$

The assumption $U_q(q(\theta),\theta,Q^e) > 0$ implies that $T_q(q(\theta)) \ge 0$.

(b) Redefining the publisher's objective function. Define the payoff function of a business with taste type θ,

$$s(\theta,Q^e) = U(q(\theta),\theta,Q^e) - T(q(\theta)). \tag{37}$$

Differentiating it with respect to θ yields

$$s_\theta(\theta,Q^e) = U_\theta(q(\theta),\theta,Q^e) + U_q(q(\theta),\theta,Q^e)q_\theta(\theta) - T_q(q(\theta)q_\theta(\theta),$$

which we combine with (35) to give

$$s_\theta(\theta,Q^e) = U_\theta(q(\theta),\theta,Q^e) > 0. \tag{38}$$

The full participation condition implies that it is optimal for the publisher to provide the lowest type businesses with zero payoff, $s(\underline{\theta},Q^e) = 0$, for any Q^e. Using it and integrating

(38) on both sides yield

$$s(\theta, Q^e) = \int_{x=\underline{\theta}}^{\theta} U_\theta(q(x), \theta, Q^e)dx, \tag{39}$$

which we substitute into (37) to obtain

$$T(q(\theta)) = U(q(\theta), \theta, Q^e) - \int_{x=\underline{\theta}}^{\theta} U_\theta(q(x), \theta, Q^e)dx. \tag{40}$$

Substituting (40) into the publisher's expected profit, we can rewrite it as

$$\max_{q(\cdot)} \int_{\underline{\theta}}^{\bar{\theta}} \Big[U(q(\theta), \theta, Q^e) - \int_{x=\underline{\theta}}^{\theta} U_\theta(q(x), \theta, Q^e)dx\Big] f(\theta|\bar{\boldsymbol{q}}^e)d\theta - C(Q^e).$$

Integrating the second part, we simplify it as

$$\max_{q(\cdot)} \int_{\underline{\theta}}^{\bar{\theta}} \Big[U(q(\theta), \theta, Q^e) - U_\theta(q(\theta), \theta, Q^e)H(\theta|\bar{\boldsymbol{q}}^e)\Big] f(\theta|\bar{\boldsymbol{q}}^e)d\theta - C(Q^e), \tag{41}$$

where $H(\theta|\bar{\boldsymbol{q}}^e) = \frac{1-F(\theta|\bar{\boldsymbol{q}}^e)}{f(\theta|\bar{\boldsymbol{q}}^e)}$ is the reciprocal of the hazard rate.

(c) Unique solution to the publisher's unconstrained problem. Combining (36) and (41), we can rewrite the (expected) publisher's profit maximization problem as

$$\max_{q(\cdot)} \int_{\underline{\theta}}^{\bar{\theta}} \Big[U(q(\theta), \theta, Q^e) - U_\theta(q(\theta), \theta, Q^e)H(\theta|\bar{\boldsymbol{q}}^e)\Big] f(\theta|\bar{\boldsymbol{q}}^e)d\theta - C(Q^e), \tag{42}$$
$$\text{s.t.} \;\; q_\theta(\theta) \geq 0.$$

If the unconstrained version of this problem has a unique solution for which $q_\theta(\theta) \geq 0$, then this is also the solution to the constrained problem. The unconstrained problem can be solved by pointwise optimizing (42) to construct $q(\cdot)$. The pointwise FOC is derived as

$$U_q(q(\theta), \theta, Q^e) - U_{q\theta}(q(\theta), \theta, Q^e)H(\theta|\bar{\boldsymbol{q}}^e) - C_Q(Q^e) = 0. \tag{43}$$

Condition (43) is also sufficient if the function

$$m(q(\theta), \theta, Q^e) = U(q(\theta), \theta, Q^e) - U_\theta(q(\theta), \theta, Q^e)H(\theta|\bar{\boldsymbol{q}}^e) \tag{44}$$

is strictly quasi-concave in q. Twice differentiating (44) with respect to q and substituting

(43) gives

$$m_{qq}(q(\theta),\theta,Q^e) = U_{qq}(q(\theta),\theta,Q^e) - U_{qq\theta}(q(\theta),\theta,Q^e)H(\theta|\bar{\boldsymbol{q}}^e) < 0, \tag{45}$$

where the inequality follows from the assumptions on $U_{qq}(q(\theta),\theta,Q^e) < 0$ and $U_{qq\theta}(q(\theta),\theta,Q^e) \geq 0$. Thus, function (44) is strictly quasi-concave in q.

(d) Confirmation of $q_\theta(\theta) > 0$. Differentiating both sides of (43) with respect to θ and rearranging it yields,

$$q_\theta(\theta) = \frac{U_{q\theta\theta}(q(\theta),\theta,Q^e)H(\theta|\bar{\boldsymbol{q}}^e) - U_{q\theta}(q(\theta),\theta,Q^e)(1 - H_\theta(\theta|\bar{\boldsymbol{q}}^e))}{U_{qq}(q(\theta),\theta,Q^e) - U_{qq\theta}(q(\theta),\theta,Q^e)H(\theta|\bar{\boldsymbol{q}}^e)}. \tag{46}$$

The denominator is strictly negative from $m(q(\theta),\theta,Q^e)$ being strictly quasi-concave. The numerator is strictly negative if $U_{q\theta\theta}(q(\theta),\theta,Q^e) \leq 0$, $U_{q\theta}(q(\theta),\theta,Q^e) > 0$, and $1-H_\theta(\theta|\bar{\boldsymbol{q}}^e) > 0$. Thus, under Assumption 1, $q_\theta(\theta) > 0$.

To sum up, under Assumption 1, given any expectations, $\bar{\boldsymbol{q}}^e$ and Q^e, the publisher's optimal price contract in equilibrium has to satisfy conditions (35) and (43), which are the necessary conditions in Proposition 1.

Step 2: Besides the necessary conditions (35) and (43), the expectations $\bar{\boldsymbol{q}}^e$ and Q^e must satisfy the fulfilled expectation conditions in equilibrium. That is, the expectations $\bar{\boldsymbol{q}}^e$ and Q^e are equal to their actual quantities in equilibrium, which are formally given by

$$q_g^* = \int_{\underline{\theta}}^{\overline{\theta}} q^*(\theta) f_j(\theta|\bar{\boldsymbol{q}}^*)d\theta, \quad \forall \;\; j \in \mathcal{J}, \tag{47}$$

$$Q^* = \int_{\underline{\theta}}^{\overline{\theta}} q^*(\theta) f(\theta|\bar{\boldsymbol{q}}^*)d\theta, \tag{48}$$

with $f_j(\theta|\bar{\boldsymbol{q}}^*)d\theta$ denoting the taste type density in industry j.

Part (ii): Part (i) has shown that given any expectations on $\bar{q}^e$ and Q^e, there exists a unique optimal price schedule satisfying (35) and (43). We now show that there exists a solution, $\bar{\boldsymbol{q}}^*$ and Q^*, satisfying (47) and (48).

Existence: Denote the vector of $\bar{\boldsymbol{q}}$ and Q by $\bar{\boldsymbol{L}} = (q_1, \cdots, q_j, Q)$. Define the fixed point mapping on $\bar{\boldsymbol{L}}$,

$$\Gamma_j(\bar{\boldsymbol{L}}) = \int_{\underline{\theta}}^{\overline{\theta}} q(\theta) f_j(\theta|\bar{\boldsymbol{L}}) d\theta, \quad \forall \quad j \in \mathcal{J},$$
$$\Gamma_{J+1}(\bar{\boldsymbol{L}}) = \int_{\underline{\theta}}^{\overline{\theta}} q(\theta) f(\theta|\bar{\boldsymbol{L}}) d\theta,$$

where $\Gamma_j(\cdot)$ is the element of the fixed-point vector $\Gamma(\bar{\boldsymbol{L}}) = (\Gamma_1(\bar{\boldsymbol{L}}), \cdots, \Gamma_J(\bar{\boldsymbol{L}}), \Gamma_{J+1}(\bar{\boldsymbol{L}}))$, $q(\theta)$ is the part of optimal contract satisfying (35) and (43) conditional on vector $\bar{\boldsymbol{L}}$.

The existence is shown in two steps. First, we show that $\Gamma_j(\bar{\boldsymbol{L}}_0) > 0$ for all j with $\bar{\boldsymbol{L}}_0 = (0, \cdots, 0, 0)$. This is satisfied from the assumption that the base intrinsic value is always positive, $V(q) > 0$. Therefore, if $\Gamma(\bar{\boldsymbol{L}})$ is bounded, then there exists a fixed point $\Gamma(\bar{\boldsymbol{L}})$ at some $\bar{\boldsymbol{L}}$. We show this in the second step. It can be shown that if $U_q(q, \theta, Q^e)$ is bounded and $U_{qq}(q, \theta, Q^e) < 0$, then $\Gamma(\bar{\boldsymbol{L}})$ is bounded. $U_{qq}(q, \theta, Q^e) < 0$ implies that the marginal utility $U_q(q, \theta, Q^e)$ is strictly decreasing in q, which together with the bounded marginal utility $U_q(q, \theta, Q^e)$, implies that the purchase $q(\theta)$ is bounded for all businesses. Therefore, $\Gamma(\bar{\boldsymbol{L}})$ is bounded. □

Proof of Lemma 1

Lemma 1. *The necessary conditions* (16) *and* (17) *in the benchmark market are equivalent to*

$$V_q(q) = \frac{T_q^{\dagger}(q)}{\overline{\theta}} \xi^{\dagger}(q) \quad \textit{and} \quad \theta^{\dagger}(q) = \frac{\overline{\theta}}{\xi^{\dagger}(q)} \tag{49}$$

for all $q \in (0, \overline{q}]$, *where*

$$\xi^{\dagger}(q) = \Big[1 - G^{\dagger}(q)\Big]^{1 - \frac{C_{Q'}(Q^{\dagger})}{T_q^{\dagger}(q)}} \exp\Big\{C_Q(Q^{\dagger}) \int_q^{\overline{q}} \frac{T_{zz}^{\dagger}(z)}{T_z^{\dagger}(z)^2} \log\Big[1 - G^{\dagger}(z)\Big] dz\Big\}, \tag{50}$$

with $\xi^{\dagger}(0) = \lim_{q \to 0} \xi^{\dagger}(q) = \overline{\theta}/\theta^{\dagger}(0)$ *and* $\xi^{\dagger}(\overline{q}) = 1$.

Proof. The proof mainly follows Equation (9) in Luo, Perrigne and Vuong (2018), which is

$$\log \frac{\theta(\alpha)}{\theta^*} = \int_0^{\alpha} \frac{1}{1-u} \left[1 - \frac{\gamma}{T_q(q(u))}\right] du, \tag{51}$$

where γ is the marginal production cost. Luo, Perrigne and Vuong (2018) show that the above equation is achieved by using the necessary conditions (16) and (17) with $W_q(\cdot,\cdot) = 0$ in the benchmark market.

First, we show Equation (50) in Lemma 1. Evaluating $\alpha = 1$ in Equation (51) gives us

$$\log \frac{\theta(1)}{\theta^*} = \int_0^1 \frac{1}{1-u} \left[1 - \frac{\gamma}{T_q(q(u))}\right] du. \tag{52}$$

Combining (51) and (52) gives

$$\log \frac{\theta(1)}{\theta(\alpha)} = \int_\alpha^1 \frac{1}{1-u} \left[1 - \frac{\gamma}{T_q(q(u))}\right] du.$$

Evaluating $\alpha = G(q)$ and implementing a change of variable $u = G(x)$, we have

$$\begin{aligned}
\log \frac{\theta(1)}{\theta(G(q))} &= \int_{G(q)}^1 \frac{1}{1-u} \left[1 - \frac{\gamma}{T_q(q(u))}\right] du \\
&= \int_q^{\overline{q}} \frac{1}{1-G(x)} \left[1 - \frac{\gamma}{T_z(z(G(x)))}\right] dG(x) \\
&= \int_q^{\overline{q}} \left[\frac{\gamma}{T_z(x)} - 1\right] d\log\left[1 - G(x)\right] \\
&= \left[\frac{\gamma}{T_q(x)} - 1\right] \log\left[1 - G(x)\right] |_q^{\overline{q}} + \gamma \int_q^{\overline{q}} \log\left[1 - G(x)\right] \frac{T_{zz}(x)}{\left[T_z(x)\right]^2} dx \\
&= \left[1 - \frac{\gamma}{T_q(q)}\right] \log\left[1 - G(q)\right] + \gamma \int_q^{\overline{q}} \log\left[1 - G(x)\right] \frac{T_{zz}(x)}{\left[T_z(x)\right]^2} dx.
\end{aligned}$$

Denoting $\gamma = C_Q(Q)$, $\theta(1) = \overline{\theta}$, and $\theta(G(q)) = \theta(q)$, defining $\xi(q) = \frac{\overline{\theta}}{\theta(q)}$, and taking the exponential of both sides of the above equation provides us with Equation (50) at the benchmark market † in Lemma 1.

Second, we show Equation (49) in Lemma 1. $\theta^\dagger(q) = \frac{\overline{\theta}}{\xi^\dagger(q)}$ in Equation (49) follows directly from the definition of $\xi(q)$. To show $V_q(q) = \frac{T_q^\dagger(q)}{\overline{\theta}} \xi^\dagger(q)$ in Equation (49), recall the necessary condition (17) at the benchmark market with $W_q(\cdot,\cdot) = 0$,

$$T_q^\dagger(q) = \theta^\dagger(q) V_q(q). \tag{53}$$

Substituting $\theta^\dagger(q) = \frac{\overline{\theta}}{\xi^\dagger(q)}$ into it gives us $V_q(q) = \frac{T_q^\dagger(q)}{\overline{\theta}} \xi^\dagger(q)$ in Equation (49).

□

2. Tables and Graphs

Table 1: Revenue Ranking by Industry Headings in SE Yellow Pages Directory

Industry heading	Revenue	Percentage
Lawyers	$ 229,036.2	4.26%
Plumbing contractors	$ 189,667.8	3.52%
Dentists	$ 131,725.8	2.45%
Personal consultants	$ 101,686.2	1.90%
Drafting services	$ 85,291.2	1.59%
Roofing contractors	$ 79,349.4	1.48%
Heating contractors	$ 76,079.4	1.41%
Self storage service	$ 74,483.4	1.38%
Electric contractors	$ 71,483.4	1.32%
Funeral homes	$ 69,765.6	1.30%
Total	$ 1,108,568.4	20.60%

Table 2: Revenue Ranking by Industry Headings in NE Yellow Pages Directory

Industry heading	Revenue	Percentage
Lawyers	$ 372,412.2	5.33%
Plumbing contractors	$ 287,662.2	4.12%
Bankruptcies-trustees	$ 136,742.8	1.96%
Dentists	$ 132,304.2	1.90%
Pest control services	$ 121,984.2	1.75%
Waterproof contractors	$ 108,791.4	1.56%
Appliances-major-sales&service	$ 99,150.0	1.42%
Roofing contractors	$ 94,075.8	1.35%
Eletric contractors	$ 91,130.4	1.31%
Rubbish removal	$ 87,197.4	1.25%
Total	$ 1,531,450.6	21.95%

Table 3: Summary Statistics in SE Yellow Pages Directory

Variable	Mean	Std. Dev.	Min	Max	Total
t	779.48	1809.27	91.20	27049.20	5.38e+6
q_{adj}	64.91	253.11	2.50	4629.68	4.48e+5
$\hat{\theta}$	0.1380	0.1297	0.0586	1.0000	-
$rent$	697.81	1903.77	4.86	18948.60	4.82e+6
$rentratio$	0.2719	0.1639	0.0506	0.5609	-

Table 4: Summary Statistics in NE Yellow Pages Directory

Variable	Mean	Std. Dev.	Min	Max	Total
t	968.27	2249.34	105.60	32371.20	6.98e+6
q_{adj}	68.48	266.86	2.43	4.626.04	4.93e+5
$\hat{\theta}$	0.1294	0.1237	0.0541	1.0000	-
$rent$	868.86	2428.13	4.44	25368.48	6.26e+6
$rentratio$	0.2698	0.1639	0.0404	0.5621	-
$netutility$	410.71	1011.94	24.01	11757.57	2.96e+6
$netratio$	0.2217	0.0040	0.2036	0.2288	-
θ^{new}	0.0973	0.1049	0.0157	0.9998	-

Table 5: Estimation of Industry-level Network Effects

Specification	I	II	II	IV
mean purchase	3.70e-04 ***	6.11e-04 ***	6.10e-04 ***	6.00e-04 ***
	(4.37e-05)	(1.08e-04)	(1.09e-04)	(1.14e-04)
mean purchase squared		-7.47e-08 *	-7.40e-08 *	-7.09e-08 *
		(3.83e-08)	(3.86e-08)	(4.01e-08)
number of firms			1.18e-05	1.24e-04
			(4.91e-05)	(2.24e-04)
number of firms squared				-1.40e-07
				(2.72e-07)
Division Fixed Effects	Yes	Yes	Yes	Yes
Adjusted R^2	0.064	0.065	0.065	0.065

Table 6: Comparisons with Different Pricing Schemes and Network Levels

	Market- and Industry-level		Industry-level		No Network	
Indicator	Linear	Nonlinear	Linear	Nonlinear	Linear	Nonlinear
Business surplus	6,978,472	0.92	5,563,156	0.83	4,344,599	0.81
Publisher profit	5,028,302	1.17	4,032,483	1.08	2,495,526	1.25
Welfare	12,006,774	1.02	9,595,639	0.94	6,840,125	0.97
Total Payment	5,884,173	1.18	4,714,625	1.09	2,853,469	1.30
Total Quantity	462,585	1.08	422,749	1.05	327,285	1.23

Table 7: Distributional Impacts of Nonlinear Pricing

	Market- and Industry-level		Industry-level		No Network	
Group and variable	Linear	Nonlinear	Linear	Nonlinear	Linear	Nonlinear
Low type						
Unit price	10.67	3.34	9.39	3.16	8.70	2.72
Surplus per business	68.08	0.40	54.48	0.34	82.01	0.09
Quantity per business	16.60	0.27	15.22	0.24	11.16	0.13
Medium type						
Unit price	10.67	2.67	9.39	2.53	8.70	2.04
Surplus per business	297.68	0.59	235.34	0.53	190.67	0.38
Quantity per business	33.00	0.39	28.87	0.39	18.28	0.63
High type						
Unit price	10.67	1.16	9.39	1.10	8.70	0.98
Surplus per business	2718.45	0.97	2168.71	0.88	1541.32	0.91
Quantity per business	152.71	1.33	140.70	1.29	107.16	1.45

Figure 1: Distribution Areas of the seven Yellow Pages Directories

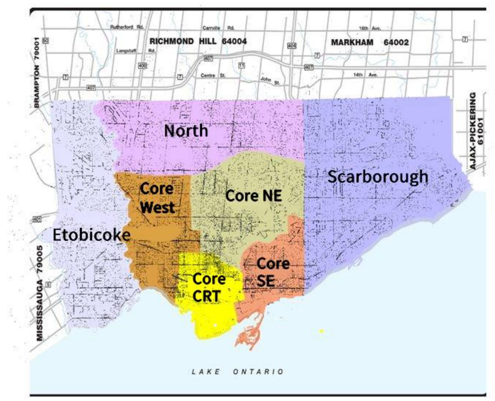

Figure 2: Estimated Intrinsic and Network Utility Functions $\hat{V}(\cdot)$, $\hat{W}^{NE}(\cdot)$

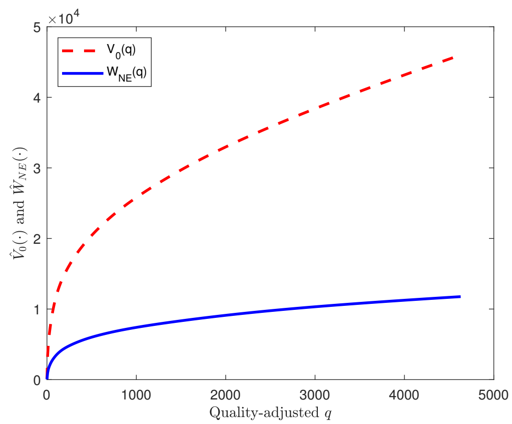

Figure 3: Estimated Type Densities $\hat{f}^{SE}(\cdot)$, $\hat{f}^{NE}(\cdot)$, $\hat{f}^{new}_{NE}(\cdot)$