NET Institute*

www.NETinst.org

**Price Competition in Two-Sided Markets
with Heterogeneous Consumers and Network Effects**

Lapo Filistrucchi and Tobias J. Klein
Tilburg University

Price Competition in Two-Sided Markets with Heterogeneous Consumers and Network Effects[*]

Lapo Filistrucchi and Tobias J. Klein[†]

October 1, 2013

Abstract

We model a two-sided market with heterogeneous customers and two heterogeneous network effects. In our model, customers on each market side care differently about both the number and the type of customers on the other side. Examples of two-sided markets are online platforms or daily newspapers. In the latter case, for instance, readership demand depends on the amount and the type of advertisements. Also, advertising demand depends on the number of readers and the distribution of readers across demographic groups. There are feedback loops because advertising demand depends on the numbers of readers, which again depends on the amount of advertising, and so on. Due to the difficulty in dealing with such feedback loops when publishers set prices on both sides of the market, most of the literature has avoided models with Bertrand competition on both sides or has resorted to simplifying assumptions such as linear demands or the presence of only one network effect. We address this issue by first presenting intuitive sufficient conditions for demand on each side to be unique given prices on both sides. We then derive sufficient conditions for the existence and uniqueness of an equilibrium in prices. For merger analysis, or any other policy simulation in the context of competition policy, it is important that equilibria exist and are unique. Otherwise, one cannot predict prices or welfare effects after a merger or a policy change. The conditions are related to the own- and cross-price effects, as well as the strength of the own and cross network effects. We show that most functional forms used in empirical work, such as logit type demand functions, tend to satisfy these conditions for realistic values of the respective parameters. Finally, using data on the Dutch daily newspaper industry, we estimate a flexible model of demand which satisfies the above conditions and evaluate the effects of a hypothetical merger and study the effects of a shrinking market for offline newspapers.

JEL Classification: L13, L40, L82.

Keywords: Two-sided markets, indirect network effects, merger simulation, equilibrium, competition policy, newspapers.

1 Introduction

Markets are called two-sided if a) firms act as platforms that sell two different products or services to two different groups of customers, b) demand of at least one group depends on demand of the other group c) firms take the inter-relatedness of demands (or indirect network effect) into account when setting prices d) customers on ones side of the market cannot pass on to customers on the other side of the market increases in the price they are asked by the platform (Rochet and Tirole, 2003; Evans, 2003; Filistrucchi, Geradin, and van Damme, 2012).

Traditional media markets are a typical example of two-sided markets (Anderson and Gabszewicz, 2008). They sell content and advertising space. Advertisers' demand for ads on a media outlet increases with the number of consumers of content (viewers, readers, listeners, etc.), while consumers of content may also be, positively or negatively, affected by the quantity of advertising. Media firms are well aware of this relationship between the two demands they face and set prices accordingly. For instance, they may lower the price on one side in order to boost demand on the other side. Free newspapers and free-to-air TV are an extreme example of such a pricing policy.

The two-sided business model is also the most common business model on the Internet. Online trading platforms, such as Amazon or eBay, or intermediaries in the advertising market, such as Google with AdSense and AdWords, sell their services to buyers and sellers that both value the popularity of the platform on the other side of the market. Two-sided is also the business model of Google, as a provider for instance of search or email services, and of Facebook as a social network. Attracting users with various free services and making advertisers pay the bill is, in fact, the same business model of free-to-air TV and free newspapers.

Not least because of the emergence of the internet, economists and policy makers have become increasingly interested in two-sided markets. In the last ten years the theoretical literature on two-sided markets has grown rapidly.[1] A key insight in this literature is that pricing decisions in two-sided markets may be very different from pricing decisions in one-sided markets. From this, it follows that analyzing a two-sided market as if it were a single-sided market may lead to mistakes and unintended consequences in the application of competition policy (Evans, 2003; Wright, 2004). For example, one may falsely predict prices to increase on both sides of the market after a merger in the absence of productive efficiency gains. On the contrary, (Chandra and Collard-Wexler, 2009) present an economic model of the newspaper

---

[1]See Rochet and Tirole (2003), Rochet and Tirole (2006) and Armstrong (2006) for the seminal theoretical studies and Rysman (2009) for a recent survey

market and show that it is not necessarily the case that a monopolist will choose to set higher prices on both market sides as compared to competing duopolists. Rather, a monopolist may choose to raise prices on one side and lower them on the other side.

Despite the growth of the theoretical literature on two-sided markets, most of the theoretical papers have either not modeled firms as setting prices on both sides of the market(Anderson and Coate, 2005) or have assumed linear demand (Armstrong, 2006) or have restricted one of the network effects to be zero (Jean J. Gabszewicz and Sonnac, 2001, 2002) or have assumed price on one side has to be zero. Similarly for the models presented in structural econometric papers: Rysman (2004) presents a model to analyze the market for phone directories in the U.S. where users of the directories clearly do not pay.[2] Kaiser and Wright (2006) limit their analysis of magazines in Germany to markets with two magazines in order to be able to apply Armstrong (2006) Hotelling duopoly model. Argentesi and Filistrucchi (2007) and Fan (forthcoming) analyze the market for daily newspapers, in Italy and in the U.S. respectively, estimating an insignificant effect of advertising on circulation and hence being able to assume no effect of advertising on readers.

In fact, while Bertrand competition is one of the standard oligopoly models used in industrial organization, the state-of-the-art in the analysis of two-sided markets does not allow to model firms as competing à la Bertrand on both sides of the market, except under the restrictive assumptions of one network effect or linear demand. This appears to be due more to technical difficulties rather than to empirical evidence showing that platforms do not set prices on both sides. In particular, there may be feedback loops in two-sided markets in the presence of two network effects. This is because, for instance, advertising demand depends on the numbers of readers, which depends on the amount of advertising, which again depends on the amount of advertising and so on. The existence of these feedback loops implies that a price increase on one side has a complex effect on both demand on that side and demand on the other side. In practice, on the one hand, it is not clear that such a loop is finite, on the other hand it may be the case that quantities on the two-sides are not unique given prices on the two-sides. As a result, a multiplicity of optimal choices in monopoly (Weyl, 2010) and a multiplicity of equilibria in oligopoly naturally arise (White and Weyl, 2012). Yet, for merger analysis, or any other policy simulation in the context of competition policy, it is important that equilibria exist and are unique. Otherwise, one cannot predict prices or welfare effects after a merger or a policy change.

---

[2]Similarly, Jeziorski (2012) analyses the market for radio in the U.S., under the reasonable assumption that listeners cannot be asked to pay even after the merger.

In general, the existence of network effects may give rise to multiple equilibria in both the consumers' coordination game and the firms' pricing game. For example, given the prices set by firms, it might be the case that all consumers choose one platform or another platform or they split among the two platforms. In fact, when consumers on one side choose a platform they choose not only based on price on that side but also based on the expected number of users on the other side and *vice versa*. In equilibrium these expectations need to be true. This is the coordination game. Uniqueness of demand given prices is a necessary condition for the existence of a unique equilibrium in the pricing game. Yet, even when demands are unique given prices, it may be the case that more than one equilibrium exists in the firm's pricing game. Clearly, if there are multiple equilibria in the coordination game, multiple equilibria in pricing game will be more likely.

To address these difficulties Weyl (2010) and White and Weyl (2012) propose to model firms as setting insulating tariffs, i.e. price schedules conditional on the quantities on the other market side, instead of setting prices. For example, an oligopolistic newspaper publisher would not set advertising prices, but advertising prices depending on circulations of its own newspaper and the rivals' newspapers. If so, White and Weyl (2012) show that there is a unique equilibrium in insulating tariffs. However, in general, Nash equilibria in pure strategies and insulating tariffs equilibria do not coincide. It is an open question to what extent conclusions regarding price or welfare effects differ qualitatively and quantitatively depending on whether firms chose prices or price schedules. More importantly, whereas there are many instances in which firms charge prices conditional on their own quantity on the other side (e.g. when a price per viewer is charged to advertisers on TV), it is instead unclear that firms actually charge prices conditional on rivals' quantities on the other side (e.g. whether the price of TV advertising on a each station changes also with the number of viewers of competing TV stations)

In this paper, we show how one can account for the feedback loops that arise when there are two network effects between the two market sides. We first derive an intuitive sufficient condition for demand to be unique given prices. This solves the issue of multiplicity of equilibria in the consumers' coordination game. We then derive sufficient conditions for the existence and uniqueness of an equilibrium in prices. Both sets of conditions are related to the strength of the own and cross network effects.

We present a general model of a two-sided market with two network effects and heterogeneous consumers. In our model, consumers on one market side care differently about the amount and the type of advertising, and advertising demand depends on both the number and the distribution of consumer demographics, such as socioeconomic status, age and gender. This has important implications for the

firms' pricing game: as the cover price increases, some readers are more likely to stop buying (say those with low incomes). At the same time, the readers who leave may be of different (lower) value to the advertisers than those who keep on buying. Moreover, if the group of readers becomes more homogeneous, advertisers will be better able to target consumers, which will result in a higher willingness to pay for advertising for a given circulation (Chandra, 2009). Therefore, equilibrium prices depend also on the relationship between price sensitivity and value to advertisers. Ignoring this may result in inaccurate predictions regarding the price in policy simulations.

Our contribution will allow competition authorities to improve their quantitative assessment of mergers in two-sided markets. In fact, when assessing a merger between two newspaper publishers, it is important to quantify the price changes on each market side, and characterize the welfare effects. This has to be done in realistic, albeit simplified, settings in which a number of newspaper publishers own more than one newspaper (Filistrucchi, Klein, and Michielsen, 2012a,b). For this, parameters of sufficiently rich demand systems need to be estimated (Berry, 1994; Berry, Levinsohn, and Pakes, 1995) and marginal costs need to be inferred from prices. Intuitively, to infer marginal costs one searches for those values of the marginal costs such that the observed prices are optimal for the firms given the demand parameter estimates (Rosse, 1970). Having done so, one can use the estimates of the marginal costs to calculate the new equilibrium prices after the merger. It is now state-of-the art among academics and practitioners to conduct such a study for one-sided markets (for example Nevo, 2000a, or Budzinski and Ruhmer, 2010). In two-sided markets, however, this has only been done for special cases of two-sided markets. For example, the econometric models of Van Cayseele and Vanormelingen (2009) and Fan (forthcoming) only model one indirect network effect. Jeziorski (2012) allows instead for two network effects, but deals with the case when one of the two prices is zero.

For merger analysis, or any other policy simulation, it is important that equilibria exist and are unique. Otherwise, one cannot predict prices or welfare effects after a merger or a policy change. In one-sided markets, both properties usually hold for the specifications that are commonly used (Vives, 2001; Mizuno, 2003). We fill a gap in the two-sided markets literature by discussing under which conditions, in a two-sided market, Nash equilibria exist and are unique.

Finally, we estimate a model with heterogeneous consumers using data on the Dutch daily newspaper industry and evaluate the effects of a hypothetical merger and of a shrinking market for offline news.

5

2 Demand

2.1 General model

There are J platforms $j = 1, \ldots, J$ that each serve two groups of customers. Demand from one group of customers depends on demand from the other group and *vice versa*. These platforms could be online platforms such as search engines or social networks, or off-line platforms such as daily newspapers or magazines. In line with our empirical application, we will henceforth think of the platforms as newspapers, but our results apply more generally to all two-sided markets in which the platform charges membership fees.

Each newspaper j is owned by a newspaper publisher f and sells advertising space to advertisers, at a price p_j^a, and subscriptions to readers, at a price p_j^r.[3] In principle, these prices can be zero or even negative on one side, in which case membership on that side is subsidized by the platforms. This is obviously the case when newspapers are distributed for free and profits are earned solely on the advertising side. Newspapers are not able to price-discriminate among the different groups of readers or advertisers.[4] This means that they charge the same subscription price to all readers and the same advertising rate to all advertisers. There are G^r demographic groups of readers. An example of a group are the high income readers between the age of 30 and 40 who live in a particular part of the country. Conversely, there are G^a groups of advertisers. Here, each group corresponds to a combination of type of advertised product (e.g. pasta or clothing) and of type of advertisement (e.g. funny or informative). Importantly, advertising demand depends not only on the total number of readers of each of the J newspapers, but also on the distribution of readers across different demographic groups. This is sensible because certain types of advertisers will be willing to pay more for advertising space if there are, say, more high income individuals who read newspaper j. Similarly, readership demand will not only depend on the amount of advertising, but also on the type of advertisements. For example, high income individuals may appreciate informative advertisements more than funny ones.[5] Figure 1 shows the effect of feedback loops if there are two

---

[3]Notice that these prices do not depend on whether advertisers and readers actually interact later on, and hence the market is a so-called two-sided non-transaction market and the prices are effectively membership fees. See also Filistrucchi, Geradin, and van Damme (2012).

[4]We here stick to the most common assumption in theoretical models of two-sided markets. For a model of price discrimination in two sided markets, see Serfes and Liu (forthcoming). In the market for daily newspapers, while the cover price is in general the same for all newspapers, also advertising list prices do not feature price discrimination. However, it may be the case that price discrimination takes place on the advertising side through the granting of personalized discounts. Since we do not observe individual discounts to advertisers, in the empirical application that follows, we maintain the assumption of no price discrimination also on the advertising side.

[5]There could also be differences in taste within a group of readers. The essential assumption we make advertising demand will only depend on the distribution of readers between groups, but not on taste differences of those readers within each
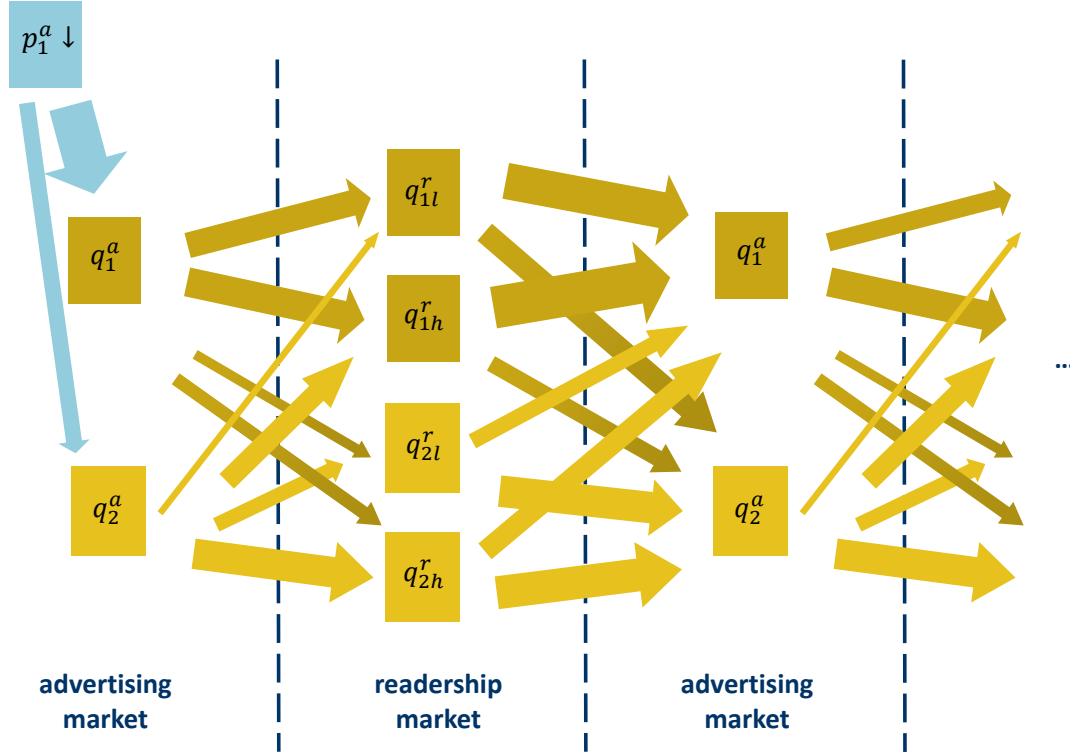
Figure 1: Feedback loops

newspapers with no groups on the advertising side, but two groups of readers, with high and low income. There are therefore two advertising demands, q_1^a and q_2^a, and four readership demands, q_{1l}^r, q_{1h}^r, q_{2l}^r and q_{2h}^r (subscripts l and h denote low and high income, respectively). Now suppose—as indicated in the upper-left corner—that the advertising price for newspaper 1 decreases. This will affect both demands on the advertising side and they will subsequently affect all four readership demands, and they will again affect all advertising demands, and so on. This process may, or may not, converge. Generally, it will converge if network effects are not too strong. We provide a sufficient condition for convergence below.

Denote the two $J \times 1$ vectors of advertising and subscription prices as p^a and p^r, respectively. More-over, for $g^a = 1, 2, ..., G^a$ denote the $J \times 1$ vector of advertising quantities of group g^a by $q_{g^a}^a$ and, for $g^r = 1, 2, ..., G^r$, denote the $J \times 1$ vector of reader quantities of group g^r by $q_{g^r}^r$. Stack them into the $G^a J \times 1$ vector of advertising quantities of all groups, $q^a = \left(q_1^{a\prime}, q_2^{a\prime}, ..., q_{G^a}^{a\prime} \right)'$, and the $G^r J \times 1$ vector

---

demographic group. That is, we assume that advertisers care about the number of readers in the high income group, but not the taste of every single member or the distribution of tastes within that group. Notice, however, that in principle one can choose a narrow definition of a group to make this assumption plausible. For empirical work, this is of course restricted by the availability of data.

7

of reader quantities $q^r = \left(q_1^{r'}, q_2^{r'}, \ldots, q_{G^a}^{r'} \right)'$. From the firms' perspective, demands at the group level on both market sides are functions of prices on the same market side and quantities at the group level on the other market side. For instance, aggregate advertising demand by one group of advertisers in newspaper j is a function of all advertising prices and the distribution of readers in demographic groups in each newspaper. These demand functions will be denoted by $q^a = q^a(p^a, q^r)$ and $q^r = q^r(p^r, q^a)$. We assume that they are continuously differentiable. It will be convenient to express demands as functions of prices only, or put differently, to work with reduced-form demand functions. We will denote these reduced-form demand functions by $q^a = \hat{q}^a(p^a, p^r)$ and $q^r = \hat{q}^r(p^a, p^r)$. In principle, quantities need not be unique for given prices (in which case these would not be functions, but correspondences). One reason for this could be a coordination problem—an issue that has received considerable attention in the theoretical literature (see, for example, Rochet and Tirole, 2003, and Armstrong, 2006). To see this, suppose that advertisers like readers and readers like advertisements. Then, it could be an equilibrium that, for given prices, all advertisers and all readers go to one newspaper. Another equilibrium could be that they all go to another newspaper. In Assumption 1 we provide a sufficient conditions for existence and uniqueness of the reduced-form demand functions given prices.[6]

Assumption 1 (network effects). *Feedback effects are not too strong in the sense that*

$$\sum_{\ell g''} \left| \sum_{kg'} \frac{\partial q_{jg}^r}{\partial q_{kg'}^a} \cdot \frac{\partial q_{kg'}^a}{\partial q_{\ell g''}^r} \right| < 1$$

and

$$\sum_{\ell g''} \left| \sum_{kg'} \frac{\partial q_{jg}^a}{\partial q_{kg'}^r} \cdot \frac{\partial q_{kg'}^r}{\partial q_{\ell g''}^a} \right| < 1$$

for all j, g, q^a, q^r.

Feedback effects are not too strong if at least one of the two network effects is not too strong. To better understand this assumption, consider the case in which advertising demand is of the constant elasticity form used in Rysman (2004),

$$\log \left(q_j^a \right) = \alpha^a + \beta^a \log \left(p_j^a \right) + \gamma^a \log \left(q_j^r \right) + \varepsilon_j.$$

Assume that readership demand is given by a standard multinomial logit model with products $j =$

$1,\ldots,J$, outside good $j = 0$ and market size M^r. Consider the simplest case in which the mean utility of a reader when purchasing good j is $\delta_j^r = \alpha^r + \beta^r p_j^r + \gamma^r q_j^a$, normalize $\delta_0^r = 0$ and denote the market shares by $s_j^r = q_j^r / M^r$. Then, the indirect network effects are

$$\frac{\partial q_j^r}{\partial q_k^a} = -M^r s_j^r s_k^r \gamma^r$$

for $j \neq \ell$ and

$$\frac{\partial q_j^r}{\partial q_j^a} = M^r s_j^r (1 - s_j^r) \gamma^r.$$

The first inequality in Assumption 1 holds if the sum of the absolute values of the changes in quantity q_j^r that originate in changes of all other quantities q_ℓ^r and affect q_j^r through q_ℓ^a is less than one. For this to be the case we need that

$$
\begin{aligned}
\sum_\ell \left| \sum_k \frac{\partial q_j^r}{\partial q_k^a} \cdot \frac{\partial q_k^a}{\partial q_\ell^r} \right| & = \sum_\ell \left| \frac{\partial q_j^r}{\partial q_\ell^a} \cdot \frac{\partial q_\ell^a}{\partial q_\ell^r} \right| \\
& = \left| M^r s_j^r (1 - s_j^r) \gamma^r \cdot \frac{\gamma^a}{q_j^r} \right| + \sum_{\ell \neq j} \left| -M^r s_j^r s_\ell^r \gamma^r \cdot \frac{\gamma^a}{q_\ell^r} \right| \\
& = \left((1 - s_j^r) - \sum_{\ell \neq j} s_j^r \right) \cdot |\gamma^a \gamma^r| \\
& = (1 - J \cdot s_j^r) \cdot |\gamma^a \gamma^r| \\
& < 1.
\end{aligned}
$$

We only have to consider $k = \ell$ in the above double sum because a change in the number of readers in newspaper ℓ will only affect advertising demand of that newspaper. This shows that the assumption restricts the absolute value of the network effect to be not too big. Observe that it always holds if one of the two network effects is zero. We will further develop the intuition underlying this restrictions in a couple of linear examples below.

Next consider another example in which demand on the advertising side is also described by a simple

logit model with parameters β^r and γ^r. Denote the indicator function by $1\{\cdot\}$. Then, Assumption 1 is

$$\sum_\ell \left| \sum_k \frac{\partial q^r_j}{\partial q^a_k} \cdot \frac{\partial q^a_k}{\partial q^r_\ell} \right| = \sum_\ell \left| \sum_k M^r \left(s^r_j 1\{j = k\} - s^r_j s^r_k \right) \gamma^r \cdot M^a \left(s^a_k 1\{k = \ell\} - s^a_k s^a_\ell \right) \gamma^a \right|$$

$$= M^a M^r \sum_\ell \left| s^r_j \sum_k \left(1\{j = k\} - s^r_k \right) \cdot \left(s^a_k 1\{k = \ell\} - s^a_k s^a_\ell \right) \gamma^a \gamma^r \right|$$

$$< 1.$$

Also here, the condition has the interpretation that the network effects are not too big. That is, given market shares on the advertising and readership side, the absolute value of $\gamma^a \gamma^r$ needs to be small enough.

Under Assumption 1 the reduced-form demand functions exist and are unique for given prices. To state this formally, stack prices into the $2J \times 1$ vector $p = (p^{a'}, p^{r'})'$, quantities into the $(G^a + G^r)J \times 1$ vector $q = (q^{a'}, q^{r'})'$, and denote the vector-valued function giving the reduced-form quantities by $\hat{q}(p)$.

Proposition 1 (existence and uniqueness of reduced-form demand functions)**.** *For any vector $p \in \mathbb{R}^{2J}$ there is a unique set of quantities $\hat{q}(p)$ if Assumption 1 holds. Moreover, for any $q_0 \in \mathbb{R}^{(G^a + G^r)J}$ the sequence of iterates $\hat{q}_0, q(p, \hat{q}_0), q(p, q(p, \hat{q}_0)), \ldots$ converges to $\hat{q}$.*

Proof. $\square$

Notice that this proposition does not say that the equilibrium of the pricing game is unique. Rather, it says that there exists a unique set of quantities for given prices. In other words, there is a unique equilibrium in the consumers' coordination game.

2.2 A linear demand example with one platform

Let us consider one newspaper facing demand for advertising and readership that is, respectively, linear in price on the same side and quantity on the other side,

$$q^a(p^a, q^r) = \alpha^a - \beta^a p^a + \gamma^a q^r \tag{1}$$
$$q^r(p^r, q^a) = \alpha^r - \beta^r p^r + \gamma^r q^a,$$

with $\beta^a, \beta^r > 0$. Solving for q^a and q^r gives

$$
\begin{aligned}
\hat{q}^a(p^a, p^r) &= \frac{1}{1 - \gamma^a \gamma^r} \cdot \{(\alpha^a + \alpha^r \gamma^a) - \beta^a p^a - \gamma^a \beta^r p^r\} \tag{2} \\
\hat{q}^r(p^a, p^r) &= \frac{1}{1 - \gamma^a \gamma^r} \cdot \{(\alpha^r + \alpha^a \gamma^r) - \gamma^r \beta^a p^a - \beta^r p^r\},
\end{aligned}
$$

provided that $\gamma^a \gamma^r \neq 1$, which is the necessary and sufficient condition for existence of the reduced-form quantities.

Alternatively, we can write (1) in matrix notation,

$$
q = \alpha + Bp + \Gamma q,
$$

with

$$
q = \begin{pmatrix} q^a \\ q^r \end{pmatrix}, p = \begin{pmatrix} p^a \\ p^r \end{pmatrix}, \alpha = \begin{pmatrix} \alpha^a \\ \alpha^r \end{pmatrix}, B = \begin{bmatrix} -\beta^a & 0 \\ 0 & -\beta^r \end{bmatrix}, \Gamma = \begin{bmatrix} 0 & \gamma^a \\ \gamma^r & 0 \end{bmatrix}.
$$

We can solve for

$$
q = (I - \Gamma)^{-1} \cdot (\alpha + Bp)
$$

provided that $\det(I - \Gamma) = 1 - \gamma^a \gamma^r \neq 0$. This shows that the above condition is actually a condition on the determinant of the matrix of network effects.

Based on (2) we can re-interpret the reduced-form demand functions as demands for complementary (if $\gamma^a, \gamma^r > 0$) or substitute (if $\gamma^a, \gamma^r < 0$) products . However, a middle case is also possible in which demand on one side depends negatively on the price of the other side but demand one the other side depends positively on the first price ($\gamma^a > 0, \gamma^r < 0$ or $\gamma^a < 0, \gamma^r > 0$). This is the case for instance if advertisers attach a higher value to newspapers with more readers but readers dislike advertising.

The condition $\gamma^a \gamma^r \neq 1$ for existence is implied by Assumption 1, which in this linear context holds whenever $|\gamma^a \gamma^r| < 1$. Formally, a solution to the above system of equations also exists if $|\gamma^a \gamma^r| > 1$. However, this solution is not meaningful in the context of demand, because the reduced-form quantities depend positively on their own price.

This shows that the conditions that guarantee existence of the reduced form demands are related to the size of the indirect network effects γ^a and γ^r. In fact, what matters is the product of the two, because

this is the module which is repeated in the loop.[7]To see this, re-write (1) as a geometric series. Define $\tilde{\alpha}^a \equiv \alpha^a - \beta^a p^a$ and $\tilde{\alpha}^r \equiv \alpha^r - \beta^r p^r$. Then,

$$
\begin{aligned}
q^a &= \tilde{\alpha}^a + \gamma^a\left(\tilde{\alpha}^r + \gamma^r\left(\tilde{\alpha}^a + \gamma^a\left(\tilde{\alpha}^r + \gamma^r\left(\ldots\right)\right)\right)\right) \\
&= \tilde{\alpha}^a + \gamma^a \gamma^r \tilde{\alpha}^a + (\gamma^a \gamma^r)^2 \tilde{\alpha}^a + \cdots + \gamma^a \tilde{\alpha}^r + \gamma^a \gamma^r \gamma^a \tilde{\alpha}^r + \gamma^a (\gamma^r \gamma^a)^2 \tilde{\alpha}^r + \ldots \\
&= (\tilde{\alpha}^a + \gamma^a \tilde{\alpha}^r)\cdot\left(1 + \gamma^a \gamma^r + (\gamma^a \gamma^r)^2 + \ldots\right)
\end{aligned}
$$

and

$$
q^r = (\tilde{\alpha}^r + \gamma^r \tilde{\alpha}^a)\cdot\left(1 + \gamma^r \gamma^a + (\gamma^r \gamma^a)^2 + \ldots\right).
$$

Both converge to the reduced form quantities (2) if the absolute value of the common ratio, $\gamma^a \gamma^r$, is less than one. This, again, is the condition in Assumption 1.

Writing demands in terms of a geometric series also shows that if one of the two network effects is zero, the reduced-form demand functions always exist because in that case the product of the two network effects is automatically zero. In that case the multiplier is equal to one.[8]

Next consider the case in $|\gamma^a \gamma^r| > 1$. We have shown above that there is a unique set of quantities in this case as well. To derive a series representation for this case re-write (1) as

$$
\begin{aligned}
\breve{q}^r(p^r, q^a) &= \frac{q^r - (\alpha^a - \beta^a p^a)}{\gamma^a} = \frac{q^r - \tilde{\alpha}^a}{\gamma^a} \\
\breve{q}^a(p^a, q^r) &= \frac{q^a - (\alpha^r - \beta^r p^r)}{\gamma^r} = \frac{q^a - \tilde{\alpha}^r}{\gamma^r}.
\end{aligned}
$$

Then, we get

$$
q^a = \left(-\frac{\tilde{\alpha}^a}{\gamma^a} - \frac{\tilde{\alpha}^r}{\gamma^a \gamma^r}\right)\cdot\left(1 + \frac{1}{\gamma^a \gamma^r} + \left(\frac{1}{\gamma^a \gamma^r}\right)^2 + \ldots\right)
$$

---

[7]One loop consists of advertising demand affecting readership demand and *thereby* affecting again advertising demand; likewise for the loop originating on the readership side.

[8]If, for instance,

$$
\begin{aligned}
q^a(p^a, q^r) &= \alpha^a - \beta^a p^a + \gamma^a q^r \\
q^r(p^r) &= \alpha^r - \beta^r p^r
\end{aligned}
$$

so that readers are not affected by advertising, then the reduced form demand functions are

$$
\begin{aligned}
\hat{q}^a(p^a, p^r) &= (\alpha^a + \alpha^r \gamma^a) - \beta^a p^a - \gamma^a \beta^r p^r \\
\hat{q}^r(p^r) &= \alpha^r - \beta^r p^r,
\end{aligned}
$$

where it appears evident that reduced form readership demand is not affected by the advertising price (because by assumption advertising quantity does not affect advertising demand), while advertising demand is affected by the cover price (since the number of readers affects demand from advertisers).

and

$$q^r = \left(-\frac{\tilde{\alpha}^r}{\gamma^r} - \frac{\tilde{\alpha}^a}{\gamma^a\gamma^r} \right) \cdot \left(1 + \frac{1}{\gamma^a\gamma^r} + \left(\frac{1}{\gamma^a\gamma^r} \right)^2 + \dots \right).$$

From these we see that indeed, quantities exist if $|\gamma^a\gamma^r| > 1$ because in that case $|1/\gamma^a\gamma^r| < 1$ and the series are in powers of $1/\gamma^a\gamma^r$. However, we have already argued above that the resulting quantities will depend positively on own prices and that therefore this case is not economically meaningful. Besides, while the functions $q^a(p^a, q^r)$ and $q^r(p^r, q^a)$ have a natural interpretation because they are primitives of the model, the functions $\check{q}^a(p^a, q^r)$ and $\check{q}^r(p^r, q^a)$ are not meaningful in the sense that, for instance, $\check{q}^a(p^a, q^r)$ means that q^a is chosen so that, for given p^a, the resulting number of readers is equal to q^r. One interpretation of Assumption 1 is therefore, that it excludes such dynamics in which there exist unique sets of quantities, but they have properties that are not economically meaningful. Here, this is because the convergent series has elements that have no economic interpretation. In the following, we will therefore only consider dynamics that satisfy Assumption 1.

2.3 A linear demand example with two platforms

Consider two newspapers facing demand functions for advertising and readership that are linear in all prices on the same side and all quantities on the other side,

$$
\begin{aligned}
q_1^a &= \alpha_1^a - \beta_{11}^a p_1^a + \beta_{12}^a p_2^a + \gamma_{11}^a q_1^r + \gamma_{21}^a q_2^r \\
q_2^a &= \alpha_2^a + \beta_{21}^a p_1^a - \beta_{22}^a p_2^a + \gamma_{21}^a q_1^r + \gamma_{22}^a q_2^r \\
q_1^r &= \alpha_1^r - \beta_{11}^r P_1^r + \beta_{21}^r p_2^r + \gamma_{11}^r q_1^a + \gamma_{12}^r q_2^a \\
q_2^r &= \alpha_2^r + \beta_{21}^r p_1^r - \beta_{22}^r p_2^r + \gamma_{21}^r q_1^a + \gamma_{22}^r q_2^a,
\end{aligned}
$$

with positive price coefficients.

This system can be written in matrix notation as

$$q = \alpha + Bp + \Gamma q,$$

with vectors

$$q = \begin{pmatrix} q^a \\ q^r \end{pmatrix} = \begin{pmatrix} q_1^a \\ q_2^a \\ q_1^r \\ q_2^r \end{pmatrix}, \alpha = \begin{pmatrix} \alpha^a \\ \alpha^r \end{pmatrix} = \begin{pmatrix} \alpha_1^a \\ \alpha_2^a \\ \alpha_1^r \\ \alpha_2^r \end{pmatrix}, p = \begin{pmatrix} p^a \\ p^r \end{pmatrix} = \begin{pmatrix} p_1^a \\ p_2^a \\ p_1^r \\ p_2^r \end{pmatrix}$$

and block-diagonal matrixes

$$\Gamma = \begin{pmatrix} 0 & \Gamma^a \\ \Gamma^r & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & \gamma_{11}^a & \gamma_{12}^a \\ 0 & 0 & \gamma_{21}^a & \gamma_{22}^a \\ \gamma_{11}^r & \gamma_{12}^r & 0 & 0 \\ \gamma_{21}^r & \gamma_{22}^r & 0 & 0 \end{pmatrix}.$$

$$B = \begin{pmatrix} B^q & 0 \\ 0 & B^r \end{pmatrix} = \begin{pmatrix} -\beta_{11}^a & \beta_{21}^a & 0 & 0 \\ \beta_{21}^a & -\beta_{22}^a & 0 & 0 \\ 0 & 0 & -\beta_{11}^r & \beta_{21}^r \\ 0 & 0 & \beta_{21}^r & -\beta_{22}^r \end{pmatrix}$$

,

We can solve for the quantities

$$q = (I - \Gamma)^{-1}(\alpha + Bp)$$

if, and only if,

$$\det (I - \Gamma) \neq 0.$$

$I - \Gamma$ is a partitioned matrix. Hence, $\det (I - \Gamma) = \det(I) \cdot \det(I - \Gamma^a \Gamma^r) = \det(I - \Gamma^a \Gamma^r) = \det(I - \Gamma^r \Gamma^a) = \det(I) \cdot \det(I - \Gamma^r \Gamma^a)$. So, the condition $\det (I - \Gamma) \neq 0$. is in fact equivalent to

$$
\begin{aligned}
\det (I - \Gamma^a \Gamma^r) &= \det \left(I - \begin{pmatrix} \gamma_{11}^a \gamma_{11}^r + \gamma_{12}^a \gamma_{21}^r & \gamma_{11}^a \gamma_{12}^r + \gamma_{12}^a \gamma_{22}^r \\ \gamma_{21}^a \gamma_{11}^r + \gamma_{22}^a \gamma_{21}^r & \gamma_{21}^a \gamma_{12}^r + \gamma_{22}^a \gamma_{22}^r \end{pmatrix} \right) \\
&= (1 - \gamma_{11}^a \gamma_{11}^r + \gamma_{12}^a \gamma_{21}^r)(1 - \gamma_{21}^a \gamma_{12}^r + \gamma_{22}^a \gamma_{22}^r) - (\gamma_{11}^a \gamma_{12}^r + \gamma_{12}^a \gamma_{22}^r)(\gamma_{21}^a \gamma_{11}^r + \gamma_{22}^a \gamma_{21}^r) \quad (3) \\
&\neq 0.
\end{aligned}
$$

Assumption 1 is that

$$\sum_{k=1}^{2}\sum_{\ell=1}^{2}|\gamma_{1k}^{a}\gamma_{k\ell}^{r}| = |\gamma_{11}^{a}\gamma_{11}^{r} + \gamma_{12}^{a}\gamma_{21}^{r}| + |\gamma_{11}^{a}\gamma_{12}^{r} + \gamma_{12}^{a}\gamma_{22}^{r}| \quad < \quad 1$$

$$\sum_{k=1}^{2}\sum_{\ell=1}^{2}|\gamma_{2k}^{a}\gamma_{k\ell}^{r}| = |\gamma_{21}^{a}\gamma_{11}^{r} + \gamma_{22}^{a}\gamma_{21}^{r}| + |\gamma_{21}^{a}\gamma_{12}^{r} + \gamma_{22}^{a}\gamma_{22}^{r}| \quad < \quad 1,$$

and implies

$$|\gamma_{11}^{a}\gamma_{11}^{r} + \gamma_{12}^{a}\gamma_{21}^{r}| \quad < \quad 1 - |\gamma_{11}^{a}\gamma_{12}^{r} + \gamma_{12}^{a}\gamma_{22}^{r}|$$

$$|\gamma_{21}^{a}\gamma_{11}^{r} + \gamma_{22}^{a}\gamma_{21}^{r}| \quad < \quad 1 - |\gamma_{21}^{a}\gamma_{12}^{r} + \gamma_{22}^{a}\gamma_{22}^{r}|.$$

Hence, the first term in parentheses in (3) is bigger than the third and the second is bigger than the fourth, and therefore $\det(I - \Gamma^{a}\Gamma^{r}) > 0$. We can thus solve for the reduced-form quantities as functions of prices only.

We can also, as in the previous example, write quantities as a geometric series,

$$q = \left(I + \Gamma\Gamma + (\Gamma\Gamma)^{2} + \dots\right) \cdot (I + \Gamma) \cdot (\alpha + Bp).$$

This series converges to

$$(1 - \Gamma\Gamma)^{-1} \cdot (I + \Gamma) \cdot (\alpha + Bp) = (1 - \Gamma)^{-1} \cdot (\alpha + Bp)$$

if the absolute value of all eigenvalues of $\Gamma\Gamma$ is strictly less than 1. Because of the block-diagonality of this matrix, with blocks $\Gamma^{a}\Gamma^{r}$ and $\Gamma^{r}\Gamma^{a}$ on the diagonal and blocks of zeros on the off-diagonal, this is the case if the absolute value of the eigenvalues of $\Gamma^{a}\Gamma^{r}$ (which are also the eigenvalues of $\Gamma^{r}\Gamma^{a}$) are strictly less than 1. They are given by

$$\lambda_{1,2} \quad = \quad \frac{(\gamma_{11}^{a}\gamma_{11}^{r} + \gamma_{12}^{a}\gamma_{21}^{r}) + (\gamma_{21}^{a}\gamma_{12}^{r} + \gamma_{22}^{a}\gamma_{22}^{r})}{2}$$
$$\pm \frac{1}{2}\sqrt{((\gamma_{11}^{a}\gamma_{11}^{r} + \gamma_{12}^{a}\gamma_{21}^{r}) + (\gamma_{21}^{a}\gamma_{12}^{r} + \gamma_{22}^{a}\gamma_{22}^{r}))^{2} - 4\det(\Gamma^{a}\Gamma^{r})}.$$

The maximal absolute value of the eigenvalues is obtained for $\det(\Gamma^a \Gamma^r) = 0$. In particular, it holds that

$$(\gamma_{11}^a \gamma_{11}^r + \gamma_{12}^a \gamma_{21}^r) = (\gamma_{21}^a \gamma_{12}^r + \gamma_{22}^a \gamma_{22}^r) = (\gamma_{11}^a \gamma_{12}^r + \gamma_{12}^a \gamma_{22}^r) = (\gamma_{21}^a \gamma_{11}^r + \gamma_{22}^a \gamma_{21}^r)$$

Then, we have that

$$\lambda_1 = (\gamma_{11}^a \gamma_{11}^r + \gamma_{12}^a \gamma_{21}^r) + (\gamma_{21}^a \gamma_{12}^r + \gamma_{22}^a \gamma_{22}^r)$$

and $\lambda_2 = 0$. Assumption 1 implies that

$$(\gamma_{11}^a \gamma_{11}^r + \gamma_{12}^a \gamma_{21}^r) = (\gamma_{21}^a \gamma_{12}^r + \gamma_{22}^a \gamma_{22}^r) = (\gamma_{11}^a \gamma_{12}^r + \gamma_{12}^a \gamma_{22}^r) = (\gamma_{21}^a \gamma_{11}^r + \gamma_{22}^a \gamma_{21}^r) < \frac{1}{2}$$

and therefore, $|\lambda_1| < 1$ and the series will converge. Gandolfo (1996, p. 117), referring to Murata (1977), shows that this holds more generally for a linear system with $J > 2$ products. Proposition 1 above can be seen as a generalization of this result to non-linear systems.

3 Competition in prices

3.1 General model

We will analyze a market in which firms compete in prices. The equilibrium concept will be Nash in pure strategies. In such an equilibrium, a firms f takes the prices of its competitors as given and maximize the sum of profits over newspapers ℓ in their portfolio $\mathscr{F}_f$,

$$\pi_f = \sum_{\ell \in \mathscr{F}_f} \left\{ (p_\ell^a - mc_\ell^a) \cdot \left(\sum_{g=1}^{G^a} q_{\ell g}^a \right) + (p_\ell^r - mc_\ell^r) \cdot \left(\sum_{g=1}^{G^r} q_{\ell g}^r \right) \right\}. \tag{4}$$

Here, quantities $q_{\ell g}^a$ and $q_{\ell g}^r$ are functions of prices on the same market side and quantities on the other market side. That is, $q_{\ell g}^a = q_{\ell g}^a(p^a, q^r)$ and $q_{\ell g}^r = q_{\ell g}^r(p^r, q^a)$. For example, advertising demand in newspaper j depends on the price of advertising in that newspaper, p_j^a, but also on the number of readers in each of the demographic groups that read that newspaper, q_{jg}^r for all $g \in G^a$, and also on all the other prices and all the other quantities on the other market side.

To formally define a Nash equilibrium in the context of our model denote the prices set by firm f by p_f. Denote the prices set by all other firms by p_{-f} and explicitly write profits as depending on prices set by f and all its competitors $-f$, $\pi(p_f, p_{-f})$.

16

Definition 1 (Nash equilibrium). A strategy profile $p^* = (p^*_f, p^*_{-f})$ is a *Nash equilibrium* if no unilateral deviation in strategy by any single firm is profitable for that firm, that is

$$\forall f, p_f : \pi_f(p^*_f, p^*_{-f}) \geq \pi_f(p_f, p^*_{-f}).$$

We will give conditions so that equilibrium prices satisfy the first order conditions. Towards this, observe that it is generally not possible to find a closed form of the first order conditions by taking the derivative of (4) with respect to the prices. This is because of the presence of feedback loops, which means that quantities on one market side depend on prices on the same market side and on quantities on the other market side. But quantities on the other market side again depend on quantities on the one market side, which again depend on prices on that first market side, and so on. However, using the reduced-form demand functions $\hat{q}^a_{jg} = \hat{q}^a_{jg}(p^a, p^r)$ and $\hat{q}^r_{jg} = \hat{q}^r_{jg}(p^a, p^r)$ we can rewrite (4) as

$$\pi_f = \sum_{\ell \in \mathscr{F}_f} \left\{ (p^a_\ell - mc^a_\ell) \cdot \left(\sum_{g=1}^{G^a} \hat{q}^a_{\ell g}(p^a, p^r) \right) + (p^r_\ell - mc^r_\ell) \cdot \left(\sum_{g=1}^{G^r} \hat{q}^r_{\ell g}(p^a, p^r) \right) \right\}$$

and the first order conditions are given by the derivative with respect to all prices,

$$\frac{\partial \pi_f}{\partial p^a_j} = q^a_j + \sum_{\ell \in \mathscr{F}_f} \left\{ (p^a_\ell - mc^a_\ell) \cdot \left(\sum_{g=1}^{G^a} \frac{\partial \hat{q}^a_{\ell g}}{\partial p^a_\ell} \right) + (p^r_\ell - mc^r_\ell) \cdot \left(\sum_{g=1}^{G^r} \frac{\partial \hat{q}^r_{\ell g}}{\partial p^a_\ell} \right) \right\} = 0 \qquad (5)$$

on the advertising side and a similar expression on the readership side.

It will be convenient to write, from now on,

$$\frac{\partial \hat{q}(p)}{\partial p'} = \left(\begin{array}{cc} \sum_{g=1}^{G^a} \partial \left[\hat{q}^a_{jg} / \partial p^{a'} \right]_j & \sum_{g=1}^{G^a} \partial \left[\hat{q}^a_{jg} / \partial p^{r'} \right]_j \\ \sum_{g=1}^{G^r} \partial \left[\hat{q}^r_{jg} / \partial p^{a'} \right]_j & \sum_{g=1}^{G^r} \partial \left[\hat{q}^r_{jg} / \partial p^{r'} \right]_j \end{array} \right).$$

The dimension of this matrix is $2J \times 2J$. $[x_j]_j$ denotes the column vector consisting of the elements x_j, stacked on top of one another in the usual way. This gives that, for example, $\left[\hat{q}^a_{jg} / \partial p^{a'} \right]_j$ is the $J \times J$ matrix of derivatives of quantities for demographic group g on the advertising side with respect to prices on the advertising side. The summation is then over demographic groups. For instance, the top-left element is the vector consisting of derivatives of market advertising demand for each of the J newspapers with respect to advertising prices. We sum over the groups of advertisers because market demand is the sum of demand by each group.

17

Denote the vector indicating which products are owned by firm f by ω_f. Then, $\Omega = \sum_f \omega_f \omega_f'$ is the Nevo (2000a, 2001)-type ownership matrix where $\Omega_{ij} = 1$ if product i and j are owned by the same company, and $\Omega_{ij} = 0$ otherwise. Define the $2J \times 2J$ matrix

$$\hat{Q} \equiv \left(\frac{\partial \hat{q}(p)}{\partial p'} \circ \begin{pmatrix} \Omega & \Omega \\ \Omega & \Omega \end{pmatrix} \right)'.$$

This is the transpose of the matrix of derivatives of the reduced form quantities, summed over demographic groups, respectively, with respect to prices, multiplied, element-wise, by the appropriate elements of the ownership matrix.

To make this approach useful in practice, Proposition 1 relates the derivatives of reduced form quantities with respect to prices to properties of the original demand functions. The latter can typically be estimated using data on quantities and prices.

Lemma 1 (price effects). *The Jacobian matrix that consists of the partial derivatives of the reduced-form demand functions $\hat{q}(p)$ with respect to the prices is given by*

$$\frac{\partial \hat{q}(p)}{\partial p'} = - \begin{pmatrix} -I & \partial q^a / \partial q^{r'} \\ \partial q^r / \partial q^{a'} & -I \end{pmatrix}^{-1} \begin{pmatrix} \partial q^a / \partial p^{a'} & 0 \\ 0 & \partial q^r / \partial p^{r'} \end{pmatrix}$$

provided that Assumption 1 holds.

Proof. See p. 44 in Appendix A. □

Observe that we can write the vector of profits earned by firm f as

$$\pi_f(p_f, p_{-f}) = \omega_f' \{ (p_a - mc_a) \circ q^a + (p_n - mc_n) \circ q^r \}.$$

and that (5) is the jth row of the system of equations

$$\hat{q}(p) + \hat{Q}(p - mc) = 0. \tag{6}$$

In total, there are $2J$ rows, one for each of the J products and 2 market sides. The first order conditions for the subscription prices are in row $J + 1$ till $2J$. From this, we get the unique vector of marginal costs that solves the first order conditions and the second order conditions.

Taking the derivative of (6) with respect to prices gives,

$$R \equiv \frac{\partial \bar{q}}{\partial p'} + (p - mc) \otimes I_J \frac{\partial \text{vec} \hat{Q}}{\partial p'} + \hat{Q},$$

where $\text{vec}A$ is the vectorizing operator that stacks the columns A on top of one another. The second order conditions are that the respective sub-matrices of R, denotes by $R_f = R(\omega_f \omega_f')$, where $R(\omega_f \omega_f')$ consists of the rows and columns of R for which the entry in $\omega_f \omega_f'$ is one, are negative definite.

A Nash equilibrium exists if there are prices such that

$$\hat{q}(p) + \hat{Q}(p - mc) = 0$$

and R_f is negative definite for all f. These are the first and second order conditions, respectively.

Conversely, if the observed quantities and prices satisfy the first order conditions because the have arisen in equilibrium, and the second order conditions hold, then one can solve for the vector of marginal costs following Rosse (1970). For this, we do not need to assume that the equilibrium is unique. However, a necessary condition is that $\hat{Q}$ is invertible. Lemma 1 shows that this condition is testable provided that demand parameters are known. In practice, one would first estimate advertising demand as a function of prices and readers, and readership demand as a function of subscription rates and the amount of advertising. Then, provided that the conditions given in Proposition (1) hold, one would use the result in Lemma (1) to calculate $\hat{Q}$. This then allows one to calculate the vector of marginal costs as

$$mc = \hat{Q}^{-1}\hat{q}(p) + p.$$

In practice, an equilibrium can be found by solving (6) numerically. This involves repeatedly calculating $\hat{q}$, following the second part of Proposition (1). In general, there could be multiple equilibria and therefore it could be of value for a policy simulation to have a set of sufficient conditions for uniqueness in hand. We provide such conditions in the following proposition. They will also imply existence and are related to properties of the firms' best reply functions.[9] First, define the best reply function for each firm f,

$$b_f(p) \equiv \arg\max_{p_f} \pi_f(p_f, p_{-f}),$$

[9]In general, these are correspondences, but here we assume that they are functions. That is, there is a single set of optimal prices for each firm, given the prices set by the competitors.

and denote the vector-valued best reply functions of all firms together by $b(p)$. Denote the jth element of this by $b_j(p)$ and the ℓth price by p_ℓ. Both can be either an advertising or a readership price of some firm.

The following proposition provides a sufficient condition for the existence of a unique equilibrium.

Proposition 2 (existence and uniqueness of equilibrium). *A unique equilibrium exists if*

$$\sum_\ell \left| \frac{\partial}{\partial p_\ell} b_j(b(p)) \right| < 1$$

at all p and for all j.

Proof. See p. in Appendix A. □

In words, the sufficient condition for existence of a unique equilibrium is that when all firms changes their prices marginally, then the sum of the additional change a firm finds it optimal to make in its jth price, after all firms, including f, have optimally reacted to the initial price change, is less than one.

Under the assumption that $\pi_f(p_f, p_{-f})$ is quasi-concave, it follows from (6) that

$$b_f(p) \equiv \left[mc - \hat{Q}^{-1} \hat{q}(p) \right]_f,$$

where $[\cdot]_f$ means that we select the rows of a vector that are related to firm f. For example, if there are two publisher and each of them owns one newspaper, and there are no groups, then the prices in p would be, in that order, the advertising price of the first firm, then the one of the second firm, then the subscription price for the first newspaper, and finally the subscription price of the second newspaper. In that case, $b(p_{-1})$ would be the first and third element of the 4×1 vector $mc - \hat{Q}^{-1} \hat{q}(p)$, because firm 1 sets the first and third price. Finally, note that here, $\hat{Q}^{-1}$ is a function of all prices, and so is $\hat{q}(p)$.

To relate the first order conditions to the condition in Proposition (2) define

$$D \equiv \frac{\partial \left(mc - \hat{Q}^{-1} \hat{q}(p) \right)}{\partial p'} = \hat{Q}^{-1} \frac{\partial \hat{Q}}{\partial p'} \hat{Q}^{-1} \hat{q}(p) - \hat{Q}^{-1} \frac{\partial \hat{q}}{\partial p'} \tag{7}$$

and denote the ijth element by $D_{i,j}$. Define a $2J$-vector Δ_ℓ with all elements equal to zero, except for the ℓth, which is equal to one. The condition is that

$$\sum_\ell \left| \sum_k D_{j,k} D_{k,\ell} \right| < 1.$$

This demonstrates that conditions for existence and uniqueness are related to the own- and cross-price effects, as well as the strength of the own and cross network effects, because D depends on the own- and cross-price effects, as well as the strength of the own and cross network effects.

There are also other, more conventional ways to establish uniqueness of equilibrium. In particular, there are three sufficient commonly employed conditions for a unique equilibrium provided that $\pi_f(p_f, p_{-f})$ is quasi-concave and differentiable in p_f for given p_{-f} (see, for instance, Vives, 2001, p. 47f). All of them can be checked numerically. The first one is a dominant diagonal property and says that the diagonal elements of

$$\partial^2 \pi_f(p_f, p_{-f})/\partial(p'_f, p'_{-f})'\partial(p'_f, p'_{-f})$$

dominate the off-diagonal elements within each row. Under this condition, one can apply the contraction mapping theorem that then implies that there is a unique fixed point of the best replies, which then is the Nash equilibrium.

The second condition is a property of the first order conditions, namely that

$$\partial^2 \pi_f(p_f, p_{-f})/\partial(p'_f, p'_{-f})'\partial(p'_f, p'_{-f})$$

is negative quasi-definite for all p_f, p_{-f}. If this Jacobian matrix is negative quasi-definite, then the Gale-Nikaido theorem implies that the map from prices to values of the first order conditions is one-to-one, which proves that there is a unique vector of prices that solves the first order conditions. Hence, the equilibrium is unique.

The third condition is that the determinant of

$$-\partial^2 \pi_f(p_f, p_{-f})/\partial(p'_f, p'_{-f})'\partial(p'_f, p'_{-f})$$

is positive whenever $\partial \pi_f(p_f, p_{-f})/\partial(p'_f, p'_{-f})' = 0$. This condition then allows one to apply the Poincaré-Hopf index theorem.

3.2   Back to the linear demand example with one platform

To better understand Proposition 2, it is instructive to go back to the linear demand example of Section 2.2 with one platform. Profits of that one platform can be written as a function of the reduced-form

demands $\hat{q}^a(p^a, p^r)$ and $\hat{q}^r(p^a, p^r)$. Denoting marginal costs by mc^a and mc^r we then have

$$
\begin{aligned}
\pi &= (p^a - mc^a) \cdot \hat{q}^a(p^a, p^r) + (p^r - mc^r) \cdot \hat{q}^r(p^a, p^r) \\
&= (p^a - mc^a) \cdot \frac{1}{1 - \gamma^a \gamma^r} \cdot \{(\alpha^a + \alpha^r \gamma^a) - \beta^a p^a - \gamma^a \beta^r p^r\} \\
&\quad + (p^r - mc^r) \cdot \frac{1}{1 - \gamma^a \gamma^r} \cdot \{(\alpha^r + \alpha^a \gamma^r) - \gamma^r \beta^a p^a - \beta^r p^r\}.
\end{aligned}
$$

where $(p^a - mc^a)$ and $(p^r - mc^r)$ are the margins on the advertising and readership side, respectively.

The first order conditions

$$
\begin{aligned}
\frac{\partial \pi}{\partial p^a} &= \hat{q}^a(p^a, p^r) + (p^a - mc^a) \cdot \frac{\partial \hat{q}^a(p^a, p^r)}{\partial p^a} + (p^r - mc^r) \cdot \frac{\partial \hat{q}^r(p^a, p^r)}{\partial p^a} = 0 \\
\frac{\partial \pi}{\partial p^r} &= \hat{q}^r(p^a, p^r) + (p^a - mc^a) \cdot \frac{\partial \hat{q}^a(p^a, p^r)}{\partial p^r} + (p^r - mc^r) \cdot \frac{\partial \hat{q}^r(p^a, p^r)}{\partial p^r} = 0
\end{aligned}
$$

are in this case

$$
\begin{aligned}
\frac{\partial \pi}{\partial p^a} &= \frac{1}{1 - \gamma^a \gamma^r} \cdot \{(\alpha^a + \alpha^r \gamma^a) - \beta^a p^a - \gamma^a \beta^r p^r\} + (p^a - mc^a) \cdot \frac{(-\beta^a)}{1 - \gamma^a \gamma^r} + (p^r - mc^r) \cdot \frac{(-\gamma^r \beta^a)}{1 - \gamma^a \gamma^r} = 0 \\
\frac{\partial \pi}{\partial p^r} &= \frac{1}{1 - \gamma^a \gamma^r} \cdot \{(\alpha^r + \alpha^a \gamma^r) - \gamma^r \beta^a p^a - \beta^r p^r\} + (p^a - mc^a) \cdot \frac{(-\gamma^a \beta^r)}{1 - \gamma^a \gamma^r} + (p^r - mc^r) \cdot \frac{(-\beta^r)}{1 - \gamma^a \gamma^r} = 0.
\end{aligned}
$$

Multiplying by $1 - \gamma^a \gamma^r \neq 0$, one obtains

$$
\{(\alpha^a + \gamma^a \alpha^r) - \beta^a p^a - \gamma^a \beta^r p^r\} + (p^a - mc^a) \cdot (-\beta^a) + (p^r - mc^r) \cdot (-\gamma^r \beta^a) = 0
$$

and

$$
\{(\gamma^r \alpha^a + \alpha^r) - \gamma^r \beta^a p^a - \beta^r p^r\} + (p^a - mc^a) \cdot (-\gamma^a \beta^r) + (p^r - mc^r) \cdot (-\beta^r) = 0,
$$

which shows that the first order conditions do not depend on the multiplier.

Solving the first equation for p^a and the second for p^r, one obtains the "internal best-reply fuctions" which give the optimal prices on each side for each price on the other side

$$p^a = \frac{(\alpha^a + \alpha^r \gamma^a)}{2\beta^a} + \frac{(\beta^a mc^a + \gamma^r \beta^a mc^r)}{2\beta^a} - \frac{(\gamma^a \beta^r + \gamma^r \beta^a)}{2\beta^a} p^r$$

$$p^r = \frac{(\alpha^r + \alpha^a \gamma^r)}{2\beta^r} + \frac{(\gamma^a \beta^r mc^a + \beta^r mc^r)}{2\beta^r} - \frac{(\gamma^r \beta^a + \gamma^a \beta^r)}{2\beta^r} p^a.$$

By a similar reasoning to the one applied in Section 2.2, an equilibrium exists and is unique under the "internal best-reply dynamics" if

$$\frac{(\gamma^r \beta^a + \gamma^a \beta^r)^2}{4\beta^a \beta^r} < 1.$$

Under the assumption of constant marginal cost, the second order conditions involve second derivatives of the profit functions with respect to prices and cross-derivatives,

$$\frac{\partial^2 \pi}{(\partial p^a)^2} = -\frac{2\beta^a}{1 - \gamma^a \gamma^r}$$

$$\frac{\partial^2 \pi}{(\partial p^r)^2} = -\frac{2\beta^r}{1 - \gamma^a \gamma^r}$$

$$\frac{\partial^2 \pi}{\partial p^a \partial p^r} = -\frac{\gamma^a \beta^r + \gamma^r \beta^a}{1 - \gamma^a \gamma^r}.$$

Strict quasi-concavity requires that the first two are negative, which holds if, and only if, $\gamma^a \gamma^r < 1$; and that the squared cross-derivative is smaller than the product of the first two second derivatives, i.e.

$$\left(\frac{\partial^2 \pi}{\partial p^a \partial p^r} \right)^2 < \frac{\partial^2 \pi}{(\partial p^a)^2} \cdot \frac{\partial^2 \pi}{(\partial p^r)^2}.$$

if and only if[10]

$$(\gamma^a \beta^r + \gamma^r \beta^a)^2 < 4\beta^a \beta^r,$$

which is equivalent to the previous condition

$$\frac{(\gamma^a \beta^r + \gamma^r \beta^a)^2}{4\beta^a \beta^r} < 1.$$

---

[10]The right hand side is always positive, and β^a and β^r on the left hand side are positive by definition. Therefore, strict quasi-concavity requires in addition to $\gamma^a \gamma^r < 1$ that neither γ^a nor γ^r are too negative.

3.3 Back to the linear demand example with two platforms

We now go back to the linear demand example of Section 2.2 with two platforms. One can distinguish two cases: a) both platforms are owned by a monopolist b) each platform is owned by a different duopolist.

Consider a monopolist owning the two platforms. Its profits can be written as a function of the reduced-form demands $\hat{q}^a(p^a, p^r)$ and $\hat{q}^r(p^a, p^r)$. These can be stacked into a vector $\hat{q}(p)$. Denoting the $J \times 1$ vectors of marginal costs by mc^a and mc^r and stacking them into the $2J \times 1$ vector mc we then have

$$\pi = (p - mc)' \cdot \hat{q}(p),$$

which, when substituting in $\hat{q} = (I - \Gamma)^{-1}(\alpha + Bp)$, becomes

$$\pi = (p - mc)' \cdot (I - \Gamma)^{-1}(\alpha + Bp).$$

The first order conditions

$$\frac{\partial \pi}{\partial p} = \hat{q}(p) + \frac{\partial \hat{q}(p)}{\partial p}' (p - mc) = 0$$

are

$$\frac{\partial \pi}{\partial p} = (I - \Gamma)^{-1}(\alpha + Bp) + \left[(I - \Gamma)^{-1}B\right]' (p - mc) = 0.$$

These can also be rewritten as

$$\frac{\partial \pi}{\partial p} = \frac{1}{det(I - \Gamma)} Adj(I - \Gamma)(\alpha + Bp) + \frac{1}{det(I - \Gamma)} \left[Adj(I - \Gamma)B\right]' (p - mc) = 0,$$

which shows that, as in the single-platform case in Section 3.2, the first order conditions do not depend on the "multiplier" $det(I - \Gamma)$.

The second order conditions are that the matrix

$$\frac{\partial \pi}{\partial p \partial p'} = \frac{\partial \left[(I - \Gamma)^{-1}(\alpha + Bp) + \left[(I - \Gamma)^{-1}B\right]' (p - mc)\right]}{\partial p'}$$

is negative semi-definite. Under the assumption that mc is constant, this is equivalent to

$$\frac{\partial \pi}{\partial p \partial p'} = (I - \Gamma)^{-1}B + \left[(I - \Gamma)^{-1}B\right]'$$

being negative semi-definite.

One can also obtain the best-reply functions

$$p = mc - \left\{\left[(I - \Gamma)^{-1}B\right]'\right\}^{-1}(I - \Gamma)^{-1}B\alpha - \left\{\left[(I - \Gamma)^{-1}B\right]'\right\}^{-1}(I - \Gamma)^{-1}Bp.$$

Taking the derivative of the best-reply functions with respect to the price vector p leads to

$$D = -\left\{\left[(I - \Gamma)^{-1}B\right]'\right\}^{-1}(I - \Gamma)^{-1}B,$$

which corresponds, in this case of a monopolist, to equation (7) above.

4   Competition between daily newspaper publishers in the Netherlands

4.1   Data

We use data from several sources. Table 1 provides an overview over the most important variables.

The top panel contains the variables we use to estimate advertising demand. They key variables here are the total amount of advertising in a given newspaper, as well as the average price paid by advertisers. The latter is calculated as the average list price that is adjusted for an average volume discount. These data are available at the quarterly level, from the fourth quarter of 1994 until the third quarter of 2009, and were obtained from The Nielsen Company (Nielsen). We follow the convention in the industry and measure the area that is covered by either content or advertisements in column millimeters. Advertising demand is also related to the number of readers and readership composition. Here, the left-out category is readers who are at least 65 years old and male readers, respectively. We therefore also use data from HOI, Instituut voor Media Auditing (HOI) on quarterly circulation at the national level, as well as data from the Print Monitor conducted by national onderzoek multimedia (NOM) on readership composition. The NOM data are available at the yearly frequency and we match the yearly observations to the other quarterly data.

The lower part of the table contains variables that we use for readership demand estimation. We define the market to be given by all individuals who are at least 13 years old and use other HOI data on

Table 1: Summary statistics

| variable	data source	level of aggregation	time span	obs.	mean	std.
advertising quantity in million column millimeter	Nielsen	newspaper-quarter	1994q4-2009q3	1,124	2.075	0.908
advertising price per column milimeter in euros of the first quarter of 2002	Nielsen	newspaper-quarter	1994q4-2009q3	1,051	3.975	3.457
circulation in million per day	HOI	newspaper-quarter	1994q4-2009q3	1,124	0.147	0.164
percentage readers age 13-34	NOM	newspaper-quarter	2000q1-2009q3	894	0.275	0.063
percentage readers age 35-49	NOM	newspaper-quarter	2000q1-2009q3	894	0.284	0.045
percentage readers age 50-64	NOM	newspaper-quarter	2000q1-2009q3	894	0.256	0.038
percentage readers female	NOM	newspaper-quarter	2000q1-2009q3	894	0.471	0.072
market share according to total circulation	HOI	newspaper-municipality-year	2002-2009	31,282	0.033	0.056
market share according to total circulation	NOM	newspaper-year	2002-2009	171	0.009	0.009
market share among individuals age 13-34	NOM	newspaper-year	2002-2009	171	0.012	0.011
market share among individuals age 35-49	NOM	newspaper-year	2002-2009	171	0.015	0.014
market share among individuals age 50-64	NOM	newspaper-year	2002-2009	171	0.011	0.010
subscription price in 2002 euros per year	Cebuco	newspaper-municipality-year	2002-2009	31,282	246.544	30.901
number individuals age 13-34	CBS	newspaper-municipality-year	2002-2009	31,282	9,913.69	18,797.45
number individuals age 35-49	CBS	newspaper-municipality-year	2002-2009	31,282	8,308.23	13,357.77
number individuals age 50-64	CBS	newspaper-municipality-year	2002-2009	31,282	6,721.45	9,347.75
number female individuals	CBS	newspaper-municipality-year	2002-2009	31,282	15,264.44	24,932.89
amount of advertising in million column milimeters	Nielsen	newspaper-municipality-year	2002-2009	31,282	7.473	3.603
total amount of content and advertising in billion column milimeters	Nielsen	newspaper-municipality-year	2002-2009	31,282	48.068	12.311
indicator for small format	Nielsen	newspaper-municipality-year	2002-2009	31,282	.253	-
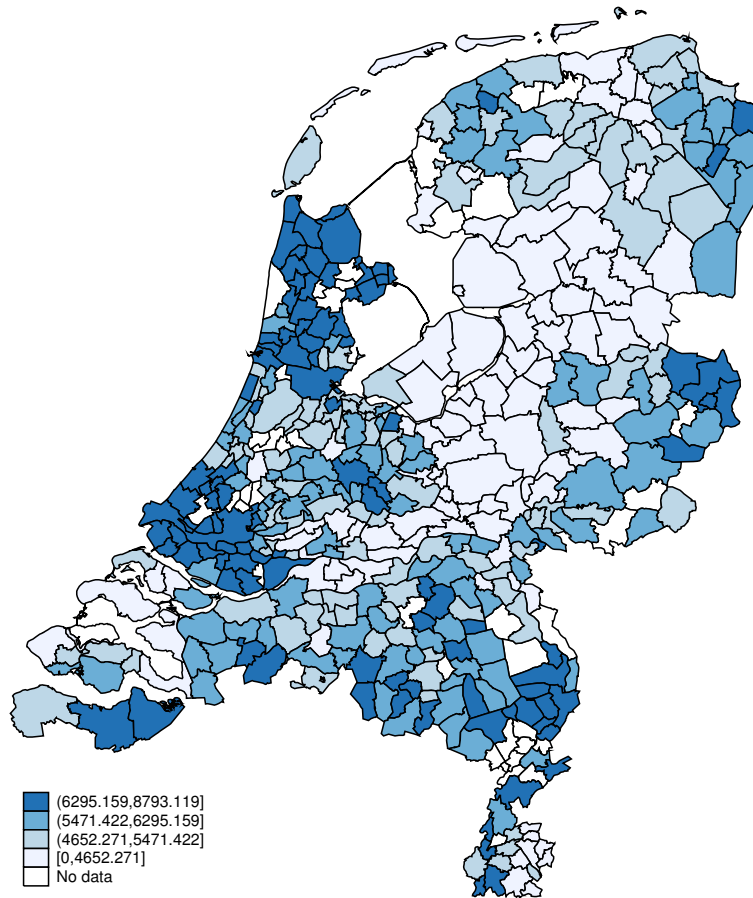
Figure 2: Readership concentration

total circulation at the municipality level and NOM data on market shares by demographic group at the national level. We use the market shares at the national level and the NOM data on the share of the people that are reached to convert reach into circulation. Data on subscription prices were obtained from Cebuco, which is related to the newspaper association NDP Nieuwsmedia. We also use data on demographics at the municipality level provided by Statistics Netherlands (CBS). For the amount of advertising, the total amount of content and the format we again use Nielsen data. A newspaper is defined to be of small format if one page contains less than 2,800 column millimeters.

Table 1 shows that the average circulation is about 150,000 and that readership is roughly equally distributed over the age categories, and gender. The market share is bigger on average when looking at the municipality level, as not all newspapers are sold in all municipalities and the 31,282 observations exclude newspapers with zero market shares in a given municipality.
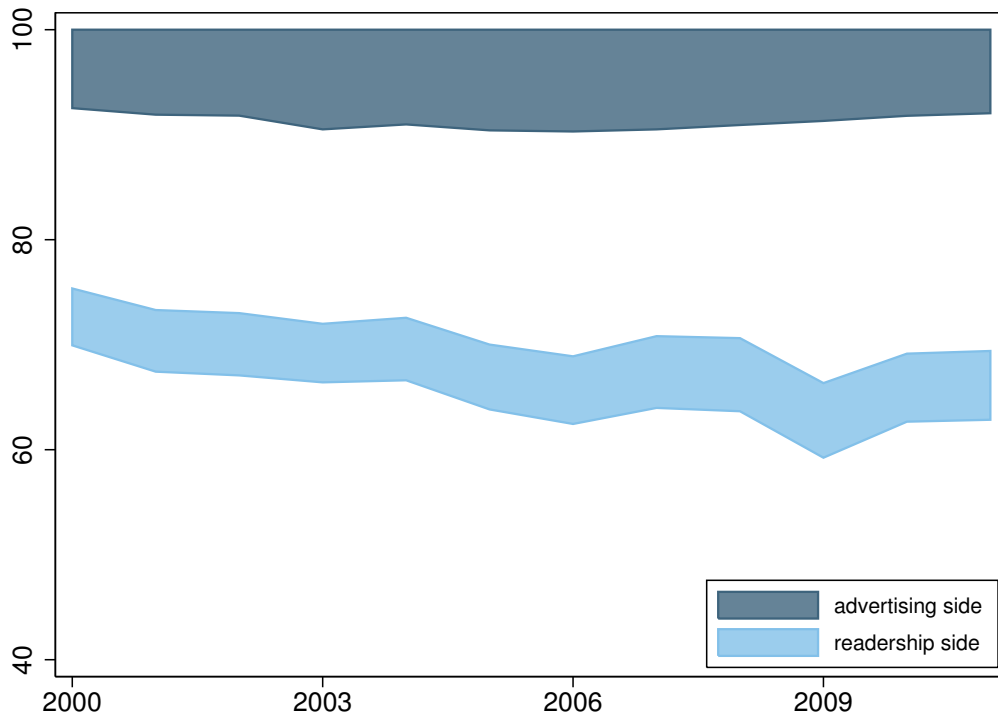
Notes: The percentage margins are calculated as follows. The starting point are 100 units of turnover. Using data on the profit relative to turnover we then calculate the cost. The short run variable cost on the readership side includes the cost of printing, which is the cost for paper, plus the cost of distribution. For the medium run cost we add the cost for maintaining technical equipment and the cost of customer care. For advertising the short run cost is zero and the medium run cost includes customer care on the advertising side. The remaining costs are counted as fixed. These include costs for the editorial office, facilities, rent and management.

Figure 3: Percentage margins over time

4.2 Descriptive evidence

The Netherlands are a small country that is extremely densely populated. The population of the U.S. is roughly 20 times bigger, while the area is 200 times bigger. Also within the country, there is considerable heterogeneity between more urban municipalities and more rural ones, both in terms of the distribution of reader characteristics and the level of competition between newspapers. Figure 2 shows a map of The Netherlands at the municipality level, in which shades of blue depict levels of the Herfindahl-Hirschman-Index (HHI) on the readership side.[11] The map shows that the level of concentration is high in the area around Amsterdam, Rotterdam and The Hague, which is in the west, and in the south. However, it is not clear whether these newspapers all operate in the same market—an implicit assumption that one has

---

[11]The HHI is defined as the sum of the squared market shares. Hence, 0 means infinitely many small firms, whereas 1 means that one firm serves the whole market. Here, we use market shares by firms and multiply multiply the obtained HHI it by $10,000$.

to make in order to think of the HHIs as measuring competition. Ultimately, this question can only be answered once we have estimated a model for readership demand, which will allow us to characterize substitution patters. Simply put, two newspapers operate in the same market if the cross-price effect is sufficiently big. This need not be the case because, for instance, most readers may want to either buy a regional level newspaper, or none at all. Or readers reading a national level tabloid newspaper may never be interested in reading a national level newspaper of high quality.

The market for daily newspapers in the Netherlands is described in Abbring and Van Ours (1994) and Filistrucchi, Klein, and Michielsen (2012a). Generally speaking, over the last decade, the market for daily newspapers has seen a downward trend in advertising volume and circulation, while both advertising prices per reader and subscription prices have increased over time.

Figure 3 shows how percentage margins on both the advertising and the readership side have changed over time. These were calculated from industry publications containing summary statistics, in particular Nederlands Uitgeversverbond (2009) for 2009 and the corresponding ones for other years. There is an area for the percentage margin, respectively, because these calculations were based on different assumptions on which costs can be seen as marginal cost. For example, for the lower bound, also costs for the hotline and customer care were counted as marginal costs on the readership side. One can see from this figure that margins were relatively stable on both market sides.

A straightforward, market-based measure for how valuable readers are to advertisers is given by the dependence of the advertising price per reader on socio-demographic characteristics of the subscribers of a newspaper. Figure 4 shows how the advertising price is related to the percentage of the readers that are between 50 and 64 years old. We have chosen this age category because on the one hand, our NOM data show that these readers are likely the ones with the highest wealth, and on the other hand we have information on age at the municipality level and are therefore able to relate an individual's price sensitivity to age. The variable that we use to measure the advertising price is the log of the advertising price per column milimeters of advertising in the newspaper and per million readers. In the figure, blue dots are for national level newspapers and red dots for regional level newspapers. We see that newspapers with more readers in the age category 50 to 64 charge higher advertising prices.

The empirical patters are also reflected in estimates of hedonic regressions that are displayed in Table 2. In all specifications, we also include a linear time trend and the percentage female readers as additional explanatory variables. In specification (2), we use wealth categories instead of the age categories and in specification (3) we use both. The robust finding is that advertising prices per reader are significantly
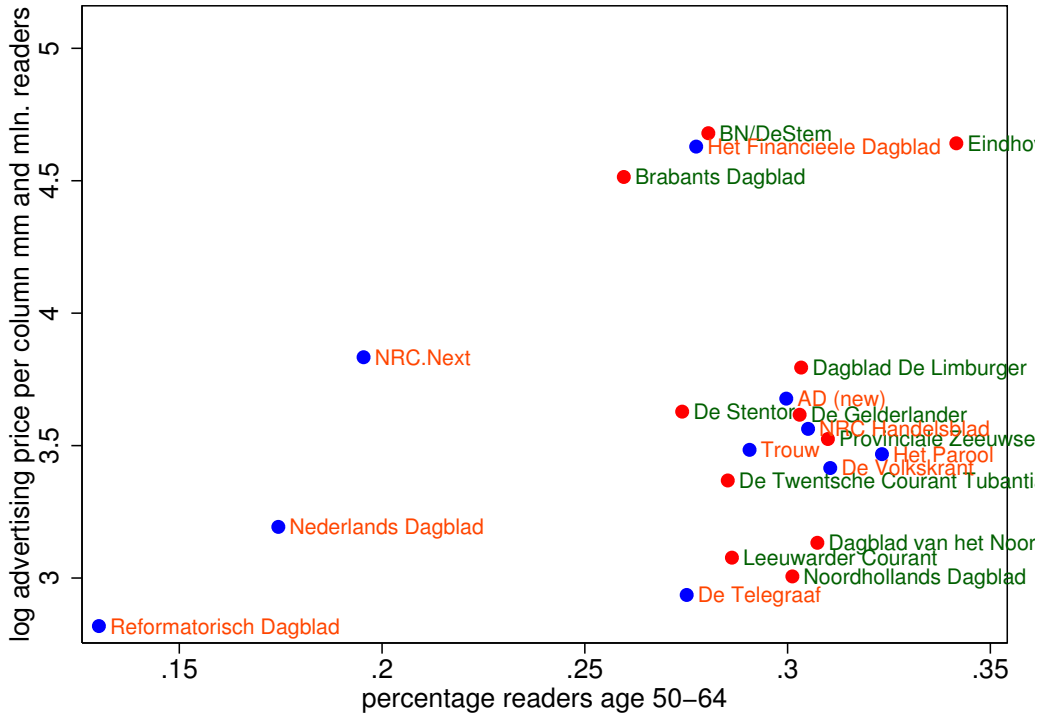
Figure 4: Relationship between advertising prices and age of readers

higher if newspapers are regional, if readers are more wealthy, and if more readers are of age 50 to 64.

4.3 Advertising side

We follow Rysman (2004) and use a constant elasticity specification,

$$\log\left(q_{jt}^a\right) = \alpha^a + \beta^a \log\left(p_{jt}^a\right) + \gamma^a \log\left(q_{jt}^r\right) + \varepsilon_{jt},$$

in which advertising demand depends only on the (own) advertising price and (own) circulation, and use the same parameter values as in Affeldt, Filistrucchi, and Klein (forthcoming). They are given in Table 3.

This advertising demand model assumes that direct cross-effects are zero on the advertising market. This is an assumption that is commonly made in this context, see also Van Cayseele and Vanormelingen (2009) and Fan (forthcoming), for instance. It means that, holding the number of subscribers constant, advertising demand in newspaper *i* depends only on the price of advertising in that newspaper, and not in others. Rysman (2004) argues that this is a reasonable assumption once readers single-home.

30

Table 2: Hedonic regressions for advertising price

	(1)	(2)	(3)
log circulation	0.840	0.784	0.797
	(0.094)	(0.070)	(0.072)
regional newspaper	0.658	0.238	0.392
	(0.178)	(0.137)	(0.152)
linear time trend	0.017	0.014	0.012
	(0.004)	(0.004)	(0.004)
percentage highest wealth category	2.168		1.230
	(0.639)		(0.538)
percentage middle wealth category	0.541		0.452
	(0.619)		(0.744)
percentage female readers	-4.630	-5.181	-4.501
	(0.806)	(0.583)	(0.511)
percentage age 35 to 49		1.248	1.033
		(1.041)	(1.120)
percentage age 50 to 64		5.491	4.451
		(1.075)	(1.137)
obs.	858	858	858

Hedonic regression with the log price per column millimeter of advertising as the dependent variable.

Table 3: Advertising demand parameters

	(1)
elasticity with respect to price (β^a)	-1.024
elasticity with respect to circulation (γ^a)	0.305

4.4 Readership side

The model for readership demand is a Berry, Levinsohn, and Pakes (1995)-type random coefficient aggregate-level mixed logit model that is estimated at the municipality level. We use data on market shares at the municipality level and data on market shares by demographic group at the national level to construct additional Petrin (2002)-type moments.[12]

Individuals i above the age of 13 in municipality m at time t choose whether or not to subscribe to one of the newspapers in the choice set $\mathscr{C}_{mt}$. The alternatives in the choice set are indexed by $j = 1, \ldots, J_{mt}$. Individuals can also choose not to subscribe to any of the newspapers, which means that they choose the

---

[12]A simpler Berry (1994)-type version of this model without random coefficients and without interactions between the price and demographics has been used Filistrucchi, Klein, and Michielsen (2012a,b) and Affeldt, Filistrucchi, and Klein (forthcoming). There, we have assumed that prices and advertising quantities are exogenous conditional on paper-region fixed effects. As we explain below, here we relax this assumption and conduct instrumental variables estimation using (close to) optimal instruments.

outside good $j = 0$. As in Lancaster (1966), alternatives are modeled as bundles of characteristics. There are observed characteristics x_{kjmt} that are collected in the vector x_{jmt}. These include the subscription price per quarter, the amount of advertising, other characteristics, as well as region-paper and year indicators. It is important to control for region-paper fixed effects (5 regions) to capture different regional focus, also for national level newspapers, and for a flexible time trend to capture the increased importance of outside options such as the evolution of the internet and the availability of free newspapers. There is also an unobserved characteristic ξ_{jmt} for each product j in each market m at each point in time t.

Utility is linear in parameters,

$$u_{ijmt} = x'_{jmt}\beta + \xi_{jmt} + \sum_{k \in \mathcal{K}} (\sigma_k v_{ki} + d'_i \pi_k) \cdot x_{kjt} + \varepsilon_{ijt}$$

for $j = 1, \ldots, J_{mt}$. For the outside good $j = 0$, $u_{i0mt} = \varepsilon_{ijt}$. $\delta_{jmt} \equiv x_{jmt}\beta + \xi_{jmt}$ is the mean utility that is enjoyed by all consumers in municipality m at time t when they choose alternative j. For alternative $j = 0$, the outside good, we have normalized that utility to zero. We denote the vector of mean utilities in municipality m at time t by δ_{mt} and the vector of mean utilities in all municipalities at time t by δ_t. Consumers tastes for attributes $k \in \mathcal{K}$ deviate from the mean valuation β according to

$$\sigma_k v_{ki} + d'_{it} \pi_k,$$

where v_{ki} is a taste shock with mean zero and variance one and d_{it} are attributes (age category and gender) for which we observe the distribution at the municipality level. Hence, $\mu_{ijt} \equiv \sum_k (\sigma_k v_{ki} + d'_i \pi_k) \cdot x_{kjt} + \varepsilon_{ijt}$ is the deviation of utility enjoyed by individual i when subscribing to newspaper j in municipality m at time t from mean utility enjoyed by the average individual. Denote the distribution function of μ_{ijt} by F_μ^{mt} and collect all parameters it depends on in the vector θ. This distribution differs across municipalities and time. Moreover, denote the distribution function of μ_{ijt} conditional on one particular demographic by $F_{\mu|d}^{mt}$. There are D such elements. In our case, there are three age groups and gender. Hence, if the second demographic is the one for the second age group, $F_{\mu|2}^{mt}$ would be the distribution of μ_{ijt} among all individuals in that age group.

Then

$$\hat{s}_{jmt}(\delta_{mt}; \theta) = \int \frac{\exp(\delta_{jmt} + \mu_{ijt})}{\sum_{j' \in \mathscr{C}_{mt}} \exp(\delta_{j'mt} + \mu_{ij't})} \, dF_\mu^{mt}(\mu_{ijt}) \tag{8}$$

is the market share of product j in municipality m at time t. We can also calculate the market share

conditional on demographic d by instead using the distribution $F_{\mu|d}^{mt}$. We will denote that market share by $\hat{s}_{jmt}^d(\delta_{mt};\theta)$.

We can aggregate unconditional market shares and market shares conditional on a vector of demographics to the national level by weighting them using the relative size of the population in municipality m at time t, conditional on the demographics in the latter case. Denote the market size in municipality m at time t by M_{mt} and the market size conditional on demographic d by M_{mt}^d. Then, the weights are given by $M_{mt}/\sum_m M_{mt}$ and $M_{mt}^d/\sum_m M_{mt}^d$, respectively. Furthermore, denote the respective implied national level market shares by

$$\hat{s}_{jt}(\delta_t;\theta) = \frac{1}{\sum_{m'} M_{mt}} \sum_m M_{mt}\hat{s}_{jmt}(\delta_t;\theta)$$

and

$$\hat{s}_{jt}^d(\delta_t;\theta) = \frac{1}{\sum_{m'} M_{mt}^d} \sum_m M_{mt}^d\hat{s}_{jmt}^d(\delta_t;\theta).$$

Observed market shares, which are data, are denoted by s_{jt} and s_{jt}^d at the national level, respectively, and s_{jmt} at the municipality level. We assume that they are observed without error.

We estimate the model by the generalized method of moments, following Berry, Levinsohn, and Pakes (1995) and Petrin (2002). Both papers implicitly or explicitly impose that implied markups are non-negative. We do so as well. Furthermore, we add a penalty if they lie outside the shaded intervals in Figure 3. Additional details on the estimation procedure are provided in Appendix B. The identifying assumption is that there are instruments that are unrelated to the demand shocks ξ_{jmt}. Throughout, we control for time and region-paper fixed effects and assume that newspaper format and the size of the newspaper are exogenous. To account for the possibility that the two key endogenous variables in the two-sided market model, price and the amount of advertising, are endogenous, we generate instruments according to the following procedure. Berry, Levinsohn, and Pakes (1995) and Reynaert and Verboven (2012) show that the optimal instruments are given by the predicted price and the predicted advertising quantity, as well as the derivative of the implied ξ_{jmt} with respect to the non-linear parameters. In order to calculate those quantities, we use an auxiliary, simplified model without random coefficients, arbitrarily set the price coefficient equal to -0.01 and the coefficient on the amount of advertising equal to 0, estimate the remaining coefficients (which are mainly time and newspaper-region fixed effects) as well as the implied ξ_{jmt}, calculate the derivative of the implied ξ_{jmt} with respect to the non-linear parameters, the implied marginal costs, and finally solve the model for prices and advertising quantities, setting all ξ_{jmt} to zero. This means that the generated prices and quantities are by construction unrelated

Table 4: Readership demand estimates

	no heterogeneity	with heterogeneity
subscription price in 2002 euros per year	-0.0168	-0.0045
	(0.0037)	(0.0096)
standard deviation random coefficient		0.0101
		(0.0003)
interacted with indicator for age 13-34		-0.0222
		(0.0009)
interacted with indicator for age 35-49		-0.0237
		(0.0009)
interacted with indicator for age 50-64		-0.0139
		(0.0007)
interacted with indicator for female		-0.0227
		(0.0007)
amount of advertising in million column millimeters	-0.0032	-0.0035
	(0.0116)	(0.0231)
standard deviation random coefficient		0.0776
		(0.0087)
total amount of content and advertising in billion column milimeters	0.0014	-0.0012
	(0.0024)	(0.0040)
indicator for small format	0.0744	0.0693
	(0.0293)	(0.0696)
region-newspaper fixed effects	yes	yes
year fixed effects	yes	yes
number observations across municipalities and time	31,282	31,282
number observations across demographics and time	0	171

to the demand shocks. Variation in prices and advertising quantities is generated from the variation in the ownership structure over time, as well as the shrinking market.

Table 4 presents demand estimates for the model without random coefficients and interactions with demographics, and for the model with random coefficients and interactions between the price and age indicators as well as an indicator for female. We also experimented with other specifications, but coefficient estimates for interactions between demographics and the amount of advertising were never significantly different from zero. Focusing on the second column for the full model, our estimates show that individuals are price-sensitive and are averse towards advertising. Both standard deviations for the random coefficients are estimated to be significantly different from zero, which means that individuals differ in their price sensitivity as well as in the sensitivity towards advertising. The left-out category for the age indicators are individuals of age 65 or older. The estimates of the coefficients on interactions between demographics and price mean that individuals between age 13 and 34 are most price sensitive and individuals between age 50 and 64 are the least price sensitive. Females are slightly more price-sensitive.

Table 5: Implications of the demand estimates

quantity	value
average own-price elasticity	-2.2411
average advertising elasticity	0.1850
average fraction of readers substituting to other newspapers	0.3743

Table 6: Equilibrium properties

quantity	value
maximal row sum of $\left\lvert \frac{\partial q^a}{\partial q^r} \cdot \frac{\partial q^r}{\partial q^{a^I}} \right\rvert$	0.0318
average implied marginal cost on the advertising side relative to price	0.1159
average implied marginal cost on the readership side relative to price	0.8623

Table 5 summarizes the implications of these estimates in terms of demand elasticities. The own-price elasticity is about -2 on average, which is similar to the one in Filistrucchi, Klein, and Michielsen (2012a) despite the fact that here, we use generated instruments. If prices are positively related to the demand shock, then one would expect that estimated price coefficients would be more negative if instruments are used that satisfy the exclusion restriction, so this points towards prices being indeed endogenous. Advertising is estimated to have a small positive effect on circulation, with an elasticity of about 0.2 on average, so that the market is found to be characterized by two indirect positive network effects between the demand for advertising and the demand for readership. Finally, the table shows the fraction of readers who would substitute towards other newspapers, rather than the outside good, when stopping to buy one particular newspaper because of a price increase. This fraction is given by the sum of the cross marginal effects, divided by the negative of the own marginal effect.

4.5 Market equilibrium

Next we relate the estimates to the theoretical results that were developed above. First, we can verify that 1 holds by calculating the maximal row sum of $\left\lvert \frac{\partial q^a}{\partial q^r} \cdot \frac{\partial q^r}{\partial q^{a^I}} \right\rvert$. The assumption holds if this quantity is less than 1, for given prices and all possible values of the quantities. We see that the value in our empirical application is much lower, at least when we calculate these derivatives locally, at the observed quantities.

The average implied marginal cost of supplying advertising is low, whereas the average implied marginal cost on the subscription side is rather high. One way to reconcile this with the percentage margins presented in Figure 3 is that newspapers may perceive some of the fixed costs as marginal when

Table 7: Merger simulation

		merging parties	non-merging parties
advertising price	percent	36.62	0.04
advertising quantity	percent	-17.72	-0.02
advertising profits	million euro	178	12
subscription price	percent	0.98	0.03
number subscriptions	percent	-2.01	0.04
subscription profits	million euro	65	39
total profits	million euro	243	51

setting the price.

4.6 Policy experiment 1: a hypothetical merger

We now study the effects of a hypothetical merger in the Dutch daily newspaper market. The hypothetical merger we investigate is between all newspapers not owned by two other publishers. The first one of those two publishers is De Persgroep, owning the Algemeen Dagblad (AD1), NRC Handelsblad (NRC), nrc.next (NRN), Het Parool (PAR), Trouw (TRO) and de Volkskant (VOL). The second one is the Telegraaf group, owning De Gooi- en Eemlander (GOO), Haarlems Dagblad (HAR), Leidsch Dagblad (LEI), Noordhollands Dagblad (NOR) and De Telegraaf (TEL). AD1 is a national-level newspaper with regional editions, NRC is a business-oriented national level newspaper, NRN is the corresponding evening edition, and PAR, TRO and VOL are other national level newspapers. The other group of newspapers consists of the regional level newspapers GOO, HAR, LEI and NOR, and the tabloid TEL. The remaining newspapers are all regional level newspapers, so what we do here is to ask the question what the effects of mergers at the regional level to one big company are.

The question whether all newspapers operate in the same market is also relevant for the policy experiments we make here because the newspapers owned by publisher 1 are mainly higher quality national level newspapers, while the newspapers owned by publisher 2 are regional level newspapers and one tabloid national level newspaper. The regional level newspapers only partly compete with the national level ones.

Table 7 summarizes the outcome of the merger simulation. The table shows that prices are predicted to increase more on the advertising side. The economic reason for this is that readers care less about advertising than advertisers care about readers, and that this is internalized by the firms when setting

36

Table 8: A shrinking market for printed newspapers

quantity	percentage change
advertising price	-0.1527
advertising quantity	0.1483
industry advertising profits	-0.0602
subscription price	0.0018
number subscriptions	-0.1378
industry subscription profits	-0.1333
industry total profits	-0.0602

prices. They will be more reluctant to increase subscription prices, because this will also have a negative effect on their profits on the advertising side, while increasing advertising prices will mostly have an impact on advertising demand as the elasticity of readership demand with respect to the amount of advertising is very small. This is a remarkable and relevant finding for competition policy, especially when a consumer—that is, reader—surplus standard is adopted. In such a case, one may argue that in a market environment in which revenues are falling the merger may be beneficial in the sense that it ensures that now merged newspaper will stay in the market, which is a benefit if "diversity of opinion" is seen as a goal, while the price for this is not paid by the readers, but mostly by the advertisers.

Overall, it turns out, the merger is not profitable for the merging parties. The competitors, however, are predicted to benefit slightly.

4.7 Policy experiment 2: shrinking market for news

Our second policy experiment is a shrinking market for printed daily newspapers. The motivation for this is the recent development that the internet as a competitor to the classic newspaper on paper has become increasingly attractive. In our model, this can be incorporated by adding a utility value to the outside option, or equivalently, subtracting that value from all utilities of the inside goods. We have changed the value of the time effect for 2009 from 0.3494 to -0.2 and have solved the model for the new equilibrium prices and quantities.

Table 8 shows that the biggest effect of this will be that subscription profits fall more than advertising profits, that the advertising price falls, but not the subscription price, and that the new equilibrium advertising quantity will be higher.

5 Concluding remarks

We propose a tractable empirical model of a two-sided market in which consumers on one market side care about the amount and the type of advertising, e.g. on an online platform or in a daily newspaper, and advertising demand depends on the distribution of consumer type such as socioeconomic status, age and gender. We show how one can account for the feedback loops that are typically present in such markets when recovering marginal costs from the first order conditions, having demand estimates in hand. Then, we derive sufficient conditions for the existence and uniqueness of an equilibrium. These conditions are related to the own- and cross-price effects, as well as the strength of the network effects. Finally, we estimate the model using data on the Dutch daily newspaper industry and evaluate the effects of a hypothetical merger.

Our results show that the conclusions may change dramatically when the two-sidedness of the market is taken into account. Our model entails the typical specification of advertising demand that is based on the idea that newspapers have a monopoly towards the advertisers when it comes to reaching their readers. Hence, the prediction of a merger simulation that ignores the two-sidedness will always be that prices will remain unaffected when firms merge. By the same token, the main and only price effects will be expected on the readership side, because there products are differentiated and firms compete in prices. However, we show that when a merger in this market is properly analyzed, which means that feedback effects between the two market sides are taken into account, then one will actually predict that prices will increase more on the advertising side, as compared to the readership side. This shows that one could make mistakes when analyzing a two-sided market as if it was a one-sided one. In this paper, we have developed general theoretical results that are directly useful to advance such structural analyzes of two sided markets with two indirect network effects and heterogeneous consumers.

References

ABBRING, J. H., AND J. C. VAN OURS (1994): "Selling news and advertising space: The economics of Dutch newspapers," *De Economist*, 142(2), 151–170.

AFFELDT, P., L. FILISTRUCCHI, AND T. J. KLEIN (forthcoming): "Upward Pricing Pressure in Two-Sided Markets," .

ANDERSON, S., AND J. GABSZEWICZ (2008): "The Media and Advertising: A Tale of Two Sided

Markets," in *Handbook of the Economics of Art and Culture*, chap. The Media and Advertising: A Tale of Two Sided Markets. Elsevier.

ANDERSON, S. P., AND S. COATE (2005): "Market provision of broadcasting: A welfare analysis," *The review of Economic studies*, 72(4), 947–972.

ARGENTESI, E., AND L. FILISTRUCCHI (2007): "Estimating Market Power in a Two-Sided Market: The Case of Newspapers," *The Journal of Applied Econometrics*, 22(7), 1247–1266.

ARMSTRONG, M. (2006): "Competition in Two-Sided Markets," *Rand Journal of Economics*, 37(3), 668–691.

BERRY, S. (1994): "Estimating Discrete-Choice Models of Product Differentiation," *Rand Journal of Economics*, 25(2), 242–262.

BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63, 841–890.

BERRY, S., J. LEVINSOHN, AND A. PAKES (2004): "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market," *Journal of Political Economy*, 112(1), 68–105.

BERRY, S., O. B. LINTON, AND A. PAKES (2004): "Limit theorems for estimating the parameters of differentiated product demand systems," *The Review of Economic Studies*, 71(3), 613–654.

BUDZINSKI, O., AND I. RUHMER (2010): "Merger simulation in competition policy: A survey," *Journal of Competition Law and Economics*, 6(2), 277–319.

CHANDRA, A. (2009): "Targeted Advertising: The Role of Subscriber Characteristics in Media Markets," *The Journal of Industrial Economics*, 57(1), 58–84.

CHANDRA, A., AND A. COLLARD-WEXLER (2009): "Mergers in Two-Sided Markets: An Application to the Canadian Newspaper Industry," *Journal of Economics and Management Strategy*, 18(4), 1045–1070.

EVANS, D. S. (2003): "The Antitrust Economics of Multi-Sided Platform Markets," *Yale Journal on Regulation*, 20, 325–382.

FAN, Y. (forthcoming): "Ownership Consolidation and Product Characteristics: A Study of the U.S. Daily Newspaper Market," *American Economic Review*.

FILISTRUCCHI, L., D. GERADIN, AND E. VAN DAMME (2012): "Identifying Two-Sided Markets," TILEC Discussion Paper No. 2012-008, Tilburg Law and Economics Center, Tilburg, The Netherlands.

FILISTRUCCHI, L., T. J. KLEIN, AND T. MICHIELSEN (2012a): "Assessing Unilateral Merger Effects in a Two-Sided Market: An Application to the Dutch Daily Newspaper Market," *Journal of Competition Law and Economics*, 8(1), 1–33.

——— (2012b): "Assessing Unilateral Merger Effects in the Daily Newspaper Market," in *Advances in the Analysis of Competition Policy and Regulation*, ed. by J. Harrington, and Y. Katsoulakos, pp. 198–213. Edward Elgar Publishing.

GANDOLFO, G. (1996): *Economic dynamics*. Springer, Berlin and New York.

HAN, C., AND P. C. PHILLIPS (2006): "GMM with many moment conditions," *Econometrica*, 74(1), 147–192.

HANSEN, L. P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50(4), 1029–1054.

JEAN J. GABSZEWICZ, D. L., AND N. SONNAC (2001): *European Economic Review* 45(4–6), 641–645.

——— (2002): "Press Advertising and the Political Differentiation of Newspapers," *Journal of Public Economic Theory*, 4(3), 317–334.

JEZIORSKI, P. (2012): "Effects of mergers in two-sided markets: Examination of the US radio industry," Mimeographed.

KAISER, U., AND J. WRIGHT (2006): "Price Structure in Two-Sided Markets: Evidence from the Magazine Industry," *International Journal of Industrial Organization*, 24, 1–28.

LANCASTER, K. J. (1966): "A New Approach to Consumer Theory," *Journal of Political Economy*, 74(2), 132–157.

MAGNUS, J. R. (2010): "On the concept of matrix derivative," *Journal of Multivariate Analysis*, 101, 2200–2206.

McFADDEN, D. (1989): "A method of simulated moments for estimation of discrete response models without numerical integration," *Econometrica*, 57(5), 995–1026.

MIZUNO, T. (2003): "On the existence of a unique price equilibrium for models of product differentiation," *International Journal of Industrial Organization*, 21(6), 761–793.

MURATA, Y. (1977): *Mathematics for Stability and Optimization of Economics Systems*. Academic Press, New York, U.S.

NEDERLANDS UITGEVERSVERBOND (2009): "De Nederlandse Dagbladpers: Jaarverslag 2009," Discussion paper, De Nederlandse Dagbladpers, Amsterdam, The Netherlands.

NEVO, A. (2000a): "Mergers with Differentiated Products: The Case of Ready-To-Eat Cereal Industry,"

Rand Journal of Economics, 31(3), 395–421.

NEVO, A. (2000b): "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand," *Journal of Economics & Management Strategy*, 9(4), 513–548.

NEVO, A. (2001): "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica*, 69(2), 307–342.

NEWEY, W., AND D. MCFADDEN (1994): "Large Sample Estimation and Hypothesis Testing," in *Handbook of Econometrics*, ed. by R. F. Engle, and D. McFadden, vol. 4, pp. 2112–2245. North Holland, Amsterdam.

PAKES, A., AND D. POLLARD (1989): "Simulation and the asymptotics of optimization estimators," *Econometrica*, 57(5), 1027–1057.

PETRIN, A. (2002): "Quantifying the Benefits of New Products: The Case of the Minivan," *Journal of Political Economy*, 110(4).

REYNAERT, M., AND F. VERBOVEN (2012): "Improving the Performance of Random Coefficients Demand Models: the Role of Optimal Instruments," Mimeographed, University of Leuven, Leuven, Belgium.

ROCHET, J.-C., AND J. TIROLE (2003): "Platform Competition in Two-Sided Markets," *Journal of the European Economic Association*, 1(4), 990–1029.

——— (2006): "Two-sided markets: a progress report," *RAND Journal of Economics*, 37(3), 645–667.

ROSSE, J. (1970): "Estimating Cost Functions Parameters without Using Cost Data: Illustrated Methodology," *Econometrica*, 38(2), 255–275.

RYSMAN, M. (2004): "Competition between Networks: A Study of the Market for Yellow Pages," *Review of Economic Studies*, 71(2), 483–512.

——— (2009): "The economics of two-sided markets," *The Journal of Economic Perspectives*, 23(3), 125–143.

SERFES, K., AND Q. LIU (forthcoming): "Price Discrimination in Two-Sided Markets," *Journal of Economics and Management Strategy*.

TRAIN, K. (2003): *Discrete Choice Methods with Simulation*. Cambridge University Press.

VAN CAYSEELE, P., AND S. VANORMELINGEN (2009): "Merger Analysis in Two-Sided Markets: the Belgian Newspaper industry," Mimeographed.

VAN DAMME, E., L. FILISTRUCCHI, D. GERADIN, S. KEUNEN, T. J. KLEIN, T. MICHIELSEN, AND J. WILEUR (2010): "Mergers in Two-Sided Markets: A Report to the NMa," Dutch Competition

Authority.

VIVES, X. (2001): *Oligopoly pricing: old ideas and new tools*. MIT press.

WEYL, G. (2010): "A Price Theory of Multi-Sided Platforms," *American Economic Review*, 100(4), 1642–1672.

WHITE, A., AND G. WEYL (2012): "Insulated Platform Competition," Unpublished Manuscript.

WRIGHT, J. (2004): "One-sided Logic in Two-sided Markets," *Review of Network Economics*, 3(1), 44–64.

A   Proofs

Proof of Proposition 1. Define the metric space $\left(\mathbb{R}_+^{(G^a+G^r)J}, d\right)$ with $d(x,y) = \|x-y\|$ being the sup-norm. This metric space is complete. Recall that we have stacked prices on both market sides into the $2J \times 1$-vector $p \equiv (p^{a'}, p^{r'})'$ and demands on both market sides and by all groups of consumers into the $(G^a+G^r)J \times 1$-vector $q \equiv (q^{a'}, q^{r'})'$. We now introduce some extra notation for this proof. In particular, denote the length of the vector of quantities by K and the demand function that gives demands for given prices and demands on the respective other market side by $\tilde{q}(p,q) \equiv (\tilde{q}^a(p^a, q^r)', \tilde{q}(p^r, q^a)')'$. In this proof, we show that under Assumption 1 $f(q) \equiv \tilde{q}(p, \tilde{q}(p,q))$, which maps quantities into quantities for given prices, is a contraction. For this, we show that there is a $\beta < 1$ such that for all $q = x, y$ in that space, $\|f(x) - f(y)\| \leq \beta \|x-y\|$.

The derivative of the jth element of this vector q with respect to the kth element is either zero—if j and k are on the same market side, or given by the indirect network effect. Define the block-diagonal matrix

$$
\Gamma(q) \equiv \begin{pmatrix} 0 & \frac{\partial \tilde{q}^a(p^a, q^r)}{\partial q^{r'}} \\ \frac{\partial \tilde{q}^r(p^r, q^a)}{\partial q^{a'}} & 0 \end{pmatrix}.
$$

Its elements are functions of the whole vector of prices and quantities, respectively. Assumption 1 is

related to the matrix

$$
\begin{aligned}
\Gamma(q)\Gamma(q) \quad &= \quad \begin{pmatrix} 0 & \frac{\partial \tilde{q}^a(p^a,q^r)}{\partial q^{r'}} \\ \frac{\partial \tilde{q}^r(p^r,q^a)}{\partial q^{a'}} & 0 \end{pmatrix} \begin{pmatrix} 0 & \frac{\partial \tilde{q}^a(p^a,q^r)}{\partial q^{r'}} \\ \frac{\partial \tilde{q}^r(p^r,q^a)}{\partial q^{a'}} & 0 \end{pmatrix} \\[2ex]
&= \quad \begin{pmatrix} \frac{\partial \tilde{q}^a(p^a,q^r)}{\partial q^{r'}} \frac{\partial \tilde{q}^r(p^r,q^a)}{\partial q^{a'}} & 0 \\ 0 & \frac{\partial \tilde{q}^a(p^a,q^r)}{\partial q^{r'}} \frac{\partial \tilde{q}^r(p^r,q^a)}{\partial q^{a'}} \end{pmatrix}.
\end{aligned}
$$

It's $j\ell$th element is

$$
-\frac{\partial \tilde{q}_j^a(p^a,q^r)}{\partial q^{r'}} \frac{\partial \tilde{q}^r(p^r,q^a)}{\partial q_\ell^{a'}} = -\sum_k \frac{\partial \tilde{q}_j^a(p^a,q^r)}{\partial q_k^{r'}} \frac{\partial \tilde{q}_k^r(p^r,q^a)}{\partial q_\ell^{a'}}
$$

if j and ℓ are both on the advertising side and

$$
-\frac{\partial \tilde{q}_j^r(p^r,q^a)}{\partial q^{a'}} \frac{\partial \tilde{q}^a(p^a,q^r)}{\partial q_\ell^{r'}} = -\sum_k \frac{\partial \tilde{q}_j^r(p^r,q^a)}{\partial q_k^{a'}} \frac{\partial \tilde{q}_k^a(p^a,q^r)}{\partial q_\ell^{r'}}
$$

if they are both on the readership side. Otherwise, they are zero.

Assumption 1 is that the sum of the absolute values of every row of $\Gamma(q)\Gamma(q)$ is less than 1, or

$$
\sum_\ell \left| \sum_k \frac{\partial \tilde{q}_j^a(p^a,q^r)}{\partial q_k^{r'}} \frac{\partial \tilde{q}_k^r(p^r,q^a)}{\partial q_\ell^{a'}} \right| < 1
$$

for every j.

By the gradient theorem of calculus we have

$$
f(x) - f(y) = \int_x^y \frac{\partial \tilde{q}(p,q)}{\partial q'} \, dq = \int_x^y \Gamma(q)\Gamma(q) \, dq,
$$

where x and y are two vectors of quantities.

Define $\lambda \equiv \|x - y\|$. The jth row of $\|f(x) - f(y)\|$ is therefore, if the jth element of q is an advertising

43

quantity,

$$
\begin{aligned}
\left| f_j(x) - f_j(y) \right| &= \left| \int_y^x \Gamma(q)\Gamma(q)\,dq \right| \\
&= \left| \int_y^x \sum_k \frac{\partial \tilde{q}_j^a(p^a, q^r)}{\partial q_k^{r'}} \frac{\partial \tilde{q}_k^r(p^r, q^a)}{\partial q^{a'}}\,dq \right| \\
&= \left| \int_y^x \sum_k \frac{\partial \tilde{q}_j^a(p^a, q^r)}{\partial q_k^{r'}} \frac{\partial \tilde{q}_k^r(p^r, q^a)}{\partial q^{a'}}\,dq \right| \\
&\leq \int_y^x \left| \sum_k \frac{\partial \tilde{q}_j^a(p^a, q^r)}{\partial q_k^{r'}} \frac{\partial \tilde{q}_k^r(p^r, q^a)}{\partial q^{a'}} \right| |dq| \\
&\leq \int_{y_1}^{x_1} \cdots \int_{y_{JG^a}}^{x_{JG^a}} \max_{q^r} \left\{ \sum_\ell \left| \sum_k \frac{\partial \tilde{q}_j^a(p^a, q^r)}{\partial q_k^{r'}} \frac{\partial \tilde{q}_k^r(p^r, q^a)}{\partial q_\ell^{a'}} \right| \right\} |dq_\ell| \\
&= \max_{q^r} \left\{ \sum_\ell \left| \sum_k \frac{\partial \tilde{q}_j^a(p^a, q^r)}{\partial q_k^{r'}} \frac{\partial \tilde{q}_k^r(p^r, q^a)}{\partial q_\ell^{a'}} \right| \right\} \int_{y_1}^{x_1} \cdots \int_{y_{JG^a}}^{x_{JG^a}} |dq_\ell| \\
&\leq \max_{q^r} \left\{ \sum_\ell \left| \sum_k \frac{\partial \tilde{q}_j^a(p^a, q^r)}{\partial q_k^{r'}} \frac{\partial \tilde{q}_k^r(p^r, q^a)}{\partial q_\ell^{a'}} \right| \right\} \int_{y_1}^{y_1 + \lambda} \cdots \int_{y_{JG^a}}^{y_{JG^a} + \lambda} |dq_\ell| \\
&= \beta \lambda
\end{aligned}
$$

with

$$
\beta = \max_{q^r} \left\{ \sum_\ell \left| \sum_k \frac{\partial \tilde{q}_j^a(p^a, q^r)}{\partial q_k^{r'}} \frac{\partial \tilde{q}_k^r(p^r, q^a)}{\partial q_\ell^{a'}} \right| \right\} < 1.
$$

The argument for the case in which the jth element of q is an advertising quantity is similar. $\square$

Proof of Lemma 1. We apply the implicit function theorem. Quantities are a function of prices and quantities on the other market side,

$$
q = q(p, q)
$$

and the total derivative of

$$
q(p, q) - q = 0
$$

is

$$
\frac{\partial (q(p,q) - q)}{\partial p'} dp + \frac{\partial (q(p,q) - q)}{\partial q'} dq = 0. \tag{9}
$$

Here, the dimension of the first Jacobian matrix is $(G^a + G^r)J \times 2J$, the dimension of dp is $2J \times 1$, the one of the second Jacobian matrix is $(G^a + G^r)J \times (G^a + G^r)J$, and the one of dq is $(G^a + G^r)J \times 1$.

Recalling that quantities depend on prices on the same market side and quantities on the other market

side we get

$$\frac{\partial\left(q(p,q)-q\right)}{\partial p'} = \begin{pmatrix} \partial q^a/\partial p^{a'} & \partial q^a/\partial p^{r'} \\ \partial q^r/\partial p^{a'} & \partial q^r/\partial p^{r'} \end{pmatrix} = \begin{pmatrix} \partial q^a/\partial p^{a'} & 0 \\ 0 & \partial q^r/\partial p^{r'} \end{pmatrix}$$

and

$$\frac{\partial\left(q(p,q)-q\right)}{\partial q'} = \begin{pmatrix} \partial q^a/\partial q^{a'} & \partial q^a/\partial q^{r'} \\ \partial q^r/\partial q^{a'} & \partial q^r/\partial q^{r'} \end{pmatrix} - \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} = \begin{pmatrix} -I & \partial q^a/\partial q^{r'} \\ \partial q^r/\partial q^{a'} & -I \end{pmatrix}.$$

Together with (9) this gives

$$\begin{pmatrix} \partial q^a/\partial p^{a'} & 0 \\ 0 & \partial q^r/\partial p^{r'} \end{pmatrix} dp + \begin{pmatrix} -I & \partial q^a/\partial q^{r'} \\ \partial q^r/\partial q^{a'} & -I \end{pmatrix} dq = 0.$$

Assumption 1 implies that

$$\begin{pmatrix} -I & \partial q^a/\partial q^{r'} \\ \partial q^r/\partial q^{a'} & -I \end{pmatrix}$$

is invertible (see Gandolfo, 1996, p. 117). Hence, the derivative of the implicit function of q as a function of p only is

$$\frac{\partial\hat{q}(p)}{\partial p'} = -\begin{pmatrix} -I & \partial q^a/\partial q^{r'} \\ \partial q^r/\partial q^{a'} & -I \end{pmatrix}^{-1} \begin{pmatrix} \partial q^a/\partial p^{a'} & 0 \\ 0 & \partial q^r/\partial p^{r'} \end{pmatrix}.$$

$\square$

Proof of Proposition 2. Define the metric space $\left(\mathbb{R}^{2J}, d\right)$ with $d(x,y) = \|x-y\|$ being the sup-norm. This metric space is complete. We show that $f(p) \equiv b(b(p))$, which maps quantities into quantities for given prices, is a contraction provided that

$$\sum_{\ell}\left|\frac{\partial}{\partial p_{\ell}}b_j\left(b(p)\right)\right| < 1.$$

For this, we show that there is a $\beta < 1$ such that for all $p = x, y$ in that space, $\|f(x) - f(y)\| \leq \beta\|x-y\|$.

By the gradient theorem of calculus we have

$$f(x) - f(y) = \int_x^y \frac{\partial b(b(p))}{\partial p'} \, dp = \int_x^y \frac{\partial b(p)}{\partial p'} \cdot \frac{\partial b(p)}{\partial p'} \, dp,$$

where x and y are two vectors of prices.

Define $\lambda \equiv \|x - y\|$. The jth row of $\|f(x) - f(y)\|$ is therefore, by arguments similar to the ones used in the proof of Proposition 1,

$$
\begin{aligned}
\left| f_j(x) - f_j(y) \right| &= \left| \int_x^y \frac{\partial b_j(p)}{\partial p'} \cdot \frac{\partial b(p)}{\partial p'} \, dp \right| \\
&\leq \int_y^x \left| \frac{\partial b_j(p)}{\partial p'} \cdot \frac{\partial b(p)}{\partial p'} \right| |dp| \\
&\leq \beta \lambda
\end{aligned}
$$

with

$$\beta = \max_p \left\{ \sum_\ell \left| \sum_k \frac{\partial b_j(p)}{\partial p_k} \cdot \frac{\partial b_k(p)}{\partial p_\ell} \right| \right\} < 1.$$

This shows that f is a contraction mapping and hence there will be a unique fixed point. $\qquad\square$

B Estimation of the readership demand model

Estimation is based on the assumption that

$$\mathbb{E}\left[(\delta_{jmt} - x'_{jmt}\beta) \cdot z_{jmt} \right] = 0 \tag{10}$$

and

$$\mathbb{E}\left[(s_{jt}^d - \hat{s}_{jt}^d(\delta_t; \theta)) \cdot z_{jt}^d \right] = 0$$

if, and only if, β is the true parameter vector. These are two sets of moment conditions, one at the product-municipality-time level, and one at the product-time level. The first set of moments is the one proposed by Berry, Levinsohn, and Pakes (1995). The second set resembles the moments in Petrin (2002) and Berry, Levinsohn, and Pakes (2004). Here, we calculate z_{jt}^d from a subset of the variables in z_{jmt} by aggregating them to the national level using weights $M_{mt}^d / \sum_m M_{mt}^d$ that sum to one. All of them are based on the idea that ξ_{jmt} is independent of the variables z_{jmt} and z_{jt}^d and the expectation is both times over products, municipalities and time. Implicitly, we assume that μ_{ijt} is independent of z_{jmt}.

We follow Berry, Levinsohn, and Pakes (1995) and Petrin (2002) and estimate the parameters using the generalized method of moments (GMM) estimator. See Berry, Linton, and Pakes (2004) for a formal discussion of the asymptotics in this context. See also Nevo (2000b) with the corresponding web appendix for additional details on the implementation. The GMM objective function is constructed by taking the following steps.

Step 1: Calculate $\hat{\delta}_{jmt}(\theta)$. Berry, Levinsohn, and Pakes (1995) show that under mild conditions we can solve for the vector of mean utilities in municipality m at time t, $\hat{\delta}_{mt}(\theta)$, with elements $\hat{\delta}_{jmt}(\theta)$. This can be done by iterating on

$$\delta_{mt}^{i+1} = \delta_{mt}^{i} + \ln(s_{jmt}) - \ln(\hat{s}_{jmt}(\delta_{mt}^{i}; \theta))$$

until convergence. This vector is a function of θ because $\hat{s}_{jmt}(\cdot; \theta)$ is a function of θ. The reason is that the distribution of μ_{ijt} in (8) depends on θ. We approximate the integral using draws of v_{ki} and the demographics. For v_{ki} we use Halton sequences of length 80 and for the demographics we draw from the observed marginal distributions of the demographics at the municipality level. We treat the number of draws as going to infinity for the asymptotics. This implies that the simulation error is negligible. See McFadden (1989), Pakes and Pollard (1989) and Train (2003) for details.

Step 2: Estimate linear parameters β. Based on the first set of moment conditions, (10), we estimate the linear parameters β using the instrumental variables two stage least squares estimator. That is, our estimates of β are given by

$$\hat{\beta} = \left(\sum_{jmt} x_{jmt} z'_{jmt} \Phi^{-1} z_{jmt} x'_{jmt} \right)^{-1} \left(\sum_{jmt} x_{jmt} z'_{jmt} \Phi^{-1} z_{jmt} \hat{\delta}_{jmt}(\theta) \right).$$

Here,

$$\Phi = \sum_{jmt} z_{jmt} z'_{jmt}$$

is the usual two stage least squares weighting matrix. It is the efficient weighting matrix if ξ_{jmt} is homoskedastic. But even if not, the estimator will be consistent.

Step 3: Calculate value of the GMM objective function for given θ. We calculate

$$g_{jmt}(\theta) \equiv (\hat{\delta}_{jmt}(\theta) - x'_{jmt} \hat{\beta}) \cdot z^-_{jmt}.$$

Here, the vector z_{jmt}^- includes only the first elements of z_{jmt}, but not the time and region-paper indicators. The reason for this is that otherwise, we would use too many moments, which would lead to the risk of obtaining biased estimates (Han and Phillips, 2006). We also stack the D sets of interactions between the deviations between observed and predicted market shares, interacted with z_{jt}^d, into the vector

$$g_{jt}(\theta) \equiv \begin{bmatrix} (s_{jt}^1 - \hat{s}_{jt}^1(\hat{\delta}_t(\theta); \theta)) \cdot z_{jt}^1 \\ \vdots \\ (s_{jt}^D - \hat{s}_{jt}^D(\hat{\delta}_t(\theta); \theta)) \cdot z_{jt}^D \end{bmatrix}.$$

Then, we expand that matrix so that each product-time combination of a national level moments is matched with the corresponding municipality level moments. That is, if there are JT observations at the national level for $g_{jt}(\theta)$ and JMT observations at the municipality level, we define the $JT \times JMT$ matrix B, in which the ijth element is one if the product time combination is the same in the ith observation at the national level as it is in the jth observation at the municipality level, and zero otherwise, and calculate

$$g_{jmt}^n(\theta) = B \cdot g_{jt}(\theta). \tag{11}$$

Expressed in this compact notation, the moment conditions are $\mathbb{E}[g_{jmt}(\theta)] = 0$ and $\mathbb{E}\left[g_{jmt}^n(\theta)\right] = 0$, if, and only if, θ is the true parameter vector. The value of the GMM objective function is

$$\begin{pmatrix} \frac{1}{JMT} \sum_{jmt} g_{jmt}(\theta) \\ \frac{1}{JMT} \sum_{jmt} g_{jmt}^n(\theta) \end{pmatrix}' W \begin{pmatrix} \frac{1}{JMT} \sum_{jmt} g_{jmt}(\theta) \\ \frac{1}{JMT} \sum_{jmt} g_{jmt}^n(\theta) \end{pmatrix},$$

where JMT is the number of observations in the data and W is a positive semi-definite weighting matrix. Summation is over products, municipalities and time.

The GMM estimator for θ_2 is then given by

$$\hat{\theta} = \arg\min_{\theta} \begin{pmatrix} \frac{1}{JMT} \sum_{jmt} g_{jmt}(\theta) \\ \frac{1}{JMT} \sum_{jmt} g_{jmt}^n(\theta) \end{pmatrix}' W \begin{pmatrix} \frac{1}{JMT} \sum_{jmt} g_{jmt}(\theta) \\ \frac{1}{JMT} \sum_{jmt} g_{jmt}^n(\theta) \end{pmatrix}.$$

The efficient estimator uses the inverse of the variance-covariance matrix of the moment conditions as the weighting matrix. In Petrin's (2002) case, this weighting matrix is block-diagonal because the two sets of moments come from two independent sampling processes and the two blocks are the respective

variance-covariance matrices of the moments. In our case, however, block-diagonality does not hold. This is because ξ_{jmt}, which enters the first set of moments, and market shares at the national level, which enter the second, are correlated. Moreover, z_{jmt} enters both sets of moments.

Hansen (1982) shows that the resulting estimator is consistent and normally distributed with variance-covariance matrix given by

$$\left(G'WG\right)^{-1}G'WVWG\left(G'WG\right)^{-1},$$

where G is the matrix of derivatives of the moments with respect to the estimated parameters, now including both β and θ. V is the variance-covariance matrix of the moment conditions,

$$V \equiv \mathbb{E}\left[\begin{pmatrix} g_{jmt}(\theta) \\ g_{jmt}^n(\theta) \end{pmatrix}\begin{pmatrix} g_{jmt}(\theta) \\ g_{jmt}^n(\theta) \end{pmatrix}'\right].$$

Given consistent, but not necessarily efficient estimates $\hat{\theta}$ we estimate V by

$$\hat{V} = \frac{1}{JMT}\sum_{jmt}\begin{pmatrix} g_{jmt}(\hat{\theta}) - \frac{1}{JMT}\sum_{jmt}g_{jmt}(\hat{\theta}) \\ g_{jt}(\hat{\theta}) - \frac{1}{JMT}\sum_{jmt}g_{jmt}^n(\hat{\theta}) \end{pmatrix}\begin{pmatrix} g_{jmt}(\hat{\theta}) - \frac{1}{JMT}\sum_{jmt}g_{jmt}(\hat{\theta}) \\ g_{jt}(\hat{\theta}) - \frac{1}{JMT}\sum_{jmt}g_{jmt}^n(\hat{\theta}) \end{pmatrix}'.$$

To obtain efficient estimates of θ we first obtain a set of initial estimates of θ and β, and estimate the variance-covariance matrix of the moment conditions. We then use its inverse as the weighting matrix W, and obtain new estimates $\hat{\theta}$. Finally, we estimate G and V and calculate the variance-covariance matrix of the estimates. In that case, the variance-covariance matrix of our estimates is given by $(G'WG)^{-1}$. See Newey and McFadden (1994) for details.

To derive G, it is useful to write the vector of moment conditions as

$$g(\beta,\theta) \equiv \mathbb{E}\begin{bmatrix} (\hat{\delta}_{jmt}(\theta) - x'_{jmt}\beta)\cdot z_{jmt}^- \\ (s_{jt}^1 - \hat{s}_{jt}^1(\hat{\delta}_t(\theta);\theta))\cdot z_{jt}^1 \\ \vdots \\ (s_{jt}^D - \hat{s}_{jt}^D(\hat{\delta}_t(\theta);\theta))\cdot z_{jt}^D \end{bmatrix}.$$

Recall that the expectation is over products, municipalities and time. Therefore, a particular combination of product and time at the national level may appear multiple times. Also, notice that we have not

substituted in $\hat{\beta}$ for β anymore. Starting from this, we have

$$
G = \left(\begin{array}{cc} \partial g(\beta,\theta)/\partial\beta' & \partial g(\beta,\theta)/\partial\theta' \end{array} \right) = \mathbb{E}\left[\left(\begin{array}{cc} -\bar{z}_{jmt}x'_{jmt} & \bar{z}_{jmt} \cdot \frac{\partial \hat{\delta}_{mt}(\theta)}{\partial\theta'} \\ -z^1_{jt} \cdot \frac{\partial \hat{s}^1_{jt}(\hat{\delta}_t(\theta);\theta)}{\partial\beta'} & -z^1_{jt} \cdot \frac{\partial \hat{s}^1_{jt}(\hat{\delta}_t(\theta);\theta)}{\partial\theta'} \\ \vdots & \vdots \\ -z^D_{jt} \cdot \frac{\partial \hat{s}^D_{jt}(\hat{\delta}_t(\theta);\theta)}{\partial\beta'} & -z^D_{jt} \cdot \frac{\partial \hat{s}^D_{jt}(\hat{\delta}_t(\theta);\theta)}{\partial\theta'} \end{array} \right) \right]
$$

$$
= \mathbb{E}\left[\begin{array}{cc} (I) & (II) \\ (III) & IV) \\ (V) & (VI) \end{array} \right].
$$

This Jacobian matrix has 6 elements, denoted by (I) through (VI). Calculating the first element is straightforward.

For (II), we have that by the implicit function theorem, the matrix with derivatives of all mean utilities $\hat{\delta}_{jmt}(\theta)$ in one market at one point in time with respect to the parameters in θ is given by

$$
\frac{\partial \hat{\delta}_{mt}(\theta)}{\partial\theta'} = -\left(\frac{\partial \hat{s}_{mt}(\hat{\delta}_{mt}(\theta);\theta)}{\partial\delta'_{mt}} \right)^{-1} \cdot \frac{\partial \hat{s}_{mt}(\delta_{mt};\theta)}{\partial\theta'}. \tag{12}
$$

To calculate the first derivative on the right hand side, define

$$
s_{ij} \equiv \frac{\exp(\delta_{jmt} + \mu_{ijt})}{\sum_{j'\in\mathscr{C}_{mt}} \exp(\delta_{j'mt} + \mu_{ij't})}.
$$

This is the integrand in (8). Then, it follows from the definition of $\hat{s}_{jmt}(\delta_{mt};\theta)$ in that equation that marginal effects are of the typical multinomial logit form,

$$
\frac{\partial \hat{s}_{mt}(\delta_{mt};\theta)}{\partial\delta'_{mt}} = \int \mathrm{dg}(s_i) - s_i s'_i \, dF^{mt}_\mu(\mu_{ijt}),
$$

where now s_i is a vector with elements s_{ij}. For the last derivative in 12, we have

$$
\frac{\partial \hat{s}_{jmt}(\delta_{mt};\theta)}{\partial\theta'} = \int s_{ij} \cdot \left(\frac{\partial \mu_{ijt}}{\partial\theta'} - \sum_{j'\in\mathscr{C}_{mt}} s_{ij'} \frac{\partial \mu_{ij't}}{\partial\theta'} \right) dF^{mt}_\mu(\mu_{ijt}).
$$

Here, elements of $\partial\mu_{mjt}/\partial\theta'$ are

$$
\frac{\partial \mu_{mjt}}{\partial\sigma_k} = v_{ki} \cdot x_{kjt}
$$

and

$$\frac{\partial \mu_{mjt}}{\partial \pi_k'} = d_i' \cdot x_{kjt}.$$

Turning to the national level moments in (III) through (VI), we have that

$$\frac{\partial \hat{s}_t^d(\hat{\delta}_t(\theta);\theta)}{\partial \beta'} = \frac{1}{\left(\sum_{m'} M_{m't}^d\right)} \cdot \sum_m M_{mt}^d \cdot \frac{\partial \hat{s}_{mt}^d(\hat{\delta}_t(\theta);\theta)}{\partial \delta_{mt}'} \cdot \frac{\partial \delta_{mt}}{\partial \beta'}$$

and

$$\frac{\partial \hat{s}_t^d(\hat{\delta}_t(\theta);\theta)}{\partial \theta'} = \frac{1}{\left(\sum_{m'} M_{m't}^d\right)} \cdot \sum_m M_{mt}^d \cdot \frac{\partial \hat{s}_{mt}^d(\hat{\delta}_t(\theta);\theta)}{\partial \theta'}.$$

These expressions involve

$$\frac{\partial \hat{s}_{mt}^d(\delta_{mt};\theta)}{\partial \delta_{mt}'} = \int \mathrm{dg}(s_i) - s_i s_i' \, dF_{\mu|d}^{mt}(\mu_{ijt})$$

and

$$\frac{\partial \delta_{jmt}}{\partial \beta'} = x_{jmt}',$$

as well as

$$\frac{\partial \hat{s}_{mt}^d(\delta_{mt};\theta)}{\partial \theta'} = \int s_{ij} \cdot \left(\frac{\partial \mu_{ijt}}{\partial \theta'} - \sum_{j' \in \mathscr{C}_{mt}} s_{ij'} \frac{\partial \mu_{ij't}}{\partial \theta'} \right) dF_{\mu|d}^{mt}(\mu_{ijt}).$$

Importantly, to compute the empirical analog of the expectation at the national level, we first need to expand these derivatives at the national level to the municipality level using the matrix B, as in (11). This then gives the estimator for G,

$$\hat{G} = \frac{1}{JMT} \sum_{jmt} \begin{pmatrix} -\bar{z}_{jmt} x_{jmt}' & \bar{z}_{jmt} \cdot \frac{\partial \hat{\delta}_{mt}(\theta)}{\partial \theta'} \\ -B_{jmt,jt} z_{jt}^1 \cdot \frac{\partial \hat{s}_{jt}^1(\hat{\delta}_t(\theta);\theta)}{\partial \beta'} & -B_{jmt,jt} z_{jt}^1 \cdot \frac{\partial \hat{s}_{jt}^1(\hat{\delta}_t(\theta);\theta)}{\partial \theta'} \\ \vdots & \vdots \\ -B_{jmt,jt} z_{jt}^D \cdot \frac{\partial \hat{s}_{jt}^D(\hat{\delta}_t(\theta);\theta)}{\partial \beta'} & -B_{jmt,jt} z_{jt}^D \cdot \frac{\partial \hat{s}_{jt}^D(\hat{\delta}_t(\theta);\theta)}{\partial \theta'} \end{pmatrix}.$$