Ownership Consolidation and Product Quality: A Study of the U.S. Daily Newspaper Market*

Ying Fan†

April 13, 2010

Abstract

I develop a structural model of newspaper markets to analyze the effects of ownership consolidation. In the model, firms choose both price and quality including the amount of non-advertising space, the number of reporters, and the number of opinion section staff. I estimate the model using a new data set on newspaper prices and characteristics. I then simulate the effect of a merger in the Minneapolis newspaper market. I also study how welfare effects of mergers vary with market characteristics. I find that ignoring adjustments of product characteristics causes substantial differences in estimated effects of mergers.

Keywords: endogenous product choice, ownership consolidation, multiple product firms, multiple discrete choice, advertising, daily newspaper market

*The previous version of the paper circulated under the title "Market Structure and Product Quality in the U.S. Daily Newspaper Market". This paper is based on various chapters of my dissertation. I am indebted to my advisors Steve Berry, Hanming Fang and Philip Haile for their continual guidance, support and encouragement. I also benefited from comments by Daniel Ackerberg, Rüdiger Bachmann, Ambarish Chandra, Matthew Gentzkow, Gautam Gowrisankaran, Justine Hastings, Alvin Klevorick, Kai-Uwe Kühn, Francine Lafontaine, Joshua Lustig, Michael Mazzeo, Sam Schulhofer-Wohl, Jesse Shapiro, Andrew Sweeting and Christopher Taber. All remaining errors are mine.

†Department of Economics, University of Michigan, 611 Tappan Street, Ann Arbor, MI 48109; yingfan@umich.edu.

1 Introduction

Do mergers affect product quality? Standard merger analyses typically study price effects only and ignore changes in the features of the product. This paper examines both. Specifically, I study the how ownership consolidation affects product characteristics and welfare in the U.S. daily newspaper market. The newspaper market provides an ideal environment for analyzing the effect of mergers on product features. First of all, individual newspapers often circulate in local markets. In contrast, many other industries (e.g., the automobile industry) provide a common set of products for the entire country. Moreover, there is substantial variation in demographics and ownership structure across these markets. This variation is crucial for this study. Secondly, like many differentiated products, the characteristics of newspapers are important for welfare. For example, after ownership consolidation, do newspaper publishers increase or decrease non-advertising space? Do they enlarge or shrink the opinion section? Do they provide more staff-written stories or utilize more material from external news agencies like the Associated Press? These questions are important for understanding welfare effects of ownership consolidation.

To address these questions, I set up a structural model of the U.S. daily newspaper market[1] that describes the demand for newspapers, the demand for advertising and publishers' decisions. The model is estimated using a new data set that I compiled, which includes information on newspaper characteristics, subscription prices, advertising rates, circulation and advertising linage for U.S. daily newspapers between 1997 and 2005.

Based on the estimates of the model parameters, I simulate the effects of a merger that two newspapers in the Minneapolis market proposed but that the Department of Justice blocked. The simulation results show that if the merger had occurred, the quality of the higher-quality newspaper would have increased while that of the other newspaper would have decreased. The increase in product differentiation would have been accompanied by a rise in both newspapers' subscription prices (by 32 dollars and 13 dollars). Overall, circulation would have declined; reader surplus would have decreased by 4.02 million dollars and publisher surplus would have increased by 15.03 million dollars. The simulation also indicates that ignoring quality adjustment leads to an underestimation of the price adjustment of one newspaper by 25 dollars per year (25% of the average annual newspaper price in the data) and an overestimation of the other newspaper's by 3 dollars. In terms of welfare effects, failure to account for quality adjustment causes an overestimation of the loss for readers by 0.02 million dollars and an underestimation of the gain for publishers by 1.07 million dollars.

[1] Other papers in the literature of newspapers are Chandra and Collard-Wexler (2009), Ferguson (1983), Genesove (1999), Gentzkow and Shapiro (2010), George (2007), Knight and Chiang (2008), Rosse (1967) and Schulhofer-Wohl and Garrido (2009). George (2007) is most closely related to this paper as she also studies market structure and product differentiation in the newspaper industry. She regresses measures of product variety on ownership concentration and finds a positive correlation between them. Since the concept of market structure is difficult to capture by a simple index, in this paper I model it explicitly.

The above case study shows how the framework provided in the paper can be used to analyze the effect of ownership consolidation for a specific market. Such a study can be computationally involved. To provide some general guidance to policy work, I then study what aspects of market characteristics are important for welfare analysis of ownership consolidation. I use the distribution of the welfare effects across markets to examine the correlation between the welfare effect of ownership consolidation and the underlying structure of the market. To this end, I quantify the welfare implications of ownership consolidation in all duopoly markets and triopoly markets. I find that readers' welfare loss is positively correlated with how much they value newspapers in general and with how important the common circulation area of the two merged parties is to these two newspapers. It is negatively correlated with the asymmetry of newspaper size measured by pre-merger circulation levels. The existence of a competitor after the merger obviously mitigates the loss in readers' welfare.

This paper contributes to the literature in several ways. First, it adds to the literature on structural merger analysis by showing that ignoring quality adjustment can be a serious omission when investigating the price effect and the welfare effect of a merger. For example, in the study of the blocked merger in the Minneapolis market, the price change of one newspaper is underestimated by 25% of the average newspaper price in the data when quality adjustment is not taken into account. This bias in price effect together with a failure to account for quality adjustment leads to a bias in the estimated welfare effect of a merger. The study of ownership consolidation in duopoly and triopoly markets also shows biases in welfare effects resulting from ignoring quality changes.

Second, this paper is related to the emerging literature on endogenous product choice, examples of which include Mazzeo (2002), Crawford and Shum (2006) and Draganska, Mazzeo and Seim (2009),[2] the last of which also studies the effect of mergers. Endogenizing product choice typically introduces important computational challenges. For example, it is generally assumed that decision makers choose characteristics in the first stage and prices in the second stage. This assumption is used to capture that the product choice decision is typically of a longer horizon. Players in the first stage take into account the impact of product choice on the equilibrium price in the second stage. However, computing equilibrium prices for each possible product choice is computationally burdensome. Papers in the literature either directly specify a profit function that is not derived from demand (such as Mazzeo (2002)) or focus on monopoly industries (such as Crawford and Shum (2006)), or examine markets with a naturally finite and discrete product choice set (such as Draganska, Mazzeo and Seim (2009)). I overcome the computational burden by using the observation that only the values of the gradient of the equilibrium price function at the data points are needed to formulate the optimality conditions for the observed product characteristics. These values are obtained from the total derivative of the first-order condition for prices. This approach

[2]Other examples in this literature include Chu (2008), Crawford and Yurukoglu (2009), Eizenberg (2009), Lustig (2009), Seim (2006), Sweeting (2007) and Watson (2009).

allows me to develop a tractable estimation routine, whereas nesting an equilibrium-solving procedure in an estimation algorithm is computationally prohibitive. My estimation strategy can be used in studying strategic behavior in a two-stage game in general.

The rest of the paper is organized as follows. Section 2 presents the structural model of newspaper markets and derives estimation equations. The data are described in Section 3. Section 4 explains the estimation approach and reports the estimation results. Section 5 studies the effect of two counterfactual ownership consolidations in the Minneapolis market. Section 6 quantifies the welfare implications of ownership consolidation in duopoly and triopoly markets and studies the correlation between the welfare effect of ownership consolidation and the underlying structure of the market. Section 7 concludes.

2 The Model

2.1 Demand

Newspaper profit comes from both selling newspapers to readers and selling advertising space to advertisers. In this section, I describe the demand for newspapers and the demand for advertising.

Since my data on newspaper circulation is at the county level, I start with the county demand for newspapers, which is derived from the aggregation of heterogeneous households' multiple discrete choices.[3,4] A multiple discrete choice model is necessary to explain duplicate readership. In the model, I set the maximum number of newspapers that a household can subscribe to at two. The model is based on Hendel (1999), although I amend it in two ways. First, I allow for decreased utility from the second choice. Second, I ensure that a household in the model buys no more than one copy of a newspaper.

Specifically, suppose all households in a county face the same choice set and the number of daily newspapers available in county c in year t is J_{ct}. A household i in this county gets utility u_{ijct} from subscribing to newspaper j in year t and utility u_{i0ct} from an outside choice.[5] The probability that household i subscribes to newspaper j is the sum of the probability that j is the first choice and that j is the second choice:

$$\Pr\left(u_{ijct} \geq \max_{h=0,\ldots,J_{ct}} u_{ihct}\right) + \sum_{j' \neq j} \Pr\left(u_{ij'ct} \geq u_{ijct} \geq \max_{h \neq j'} u_{ihct}, \;\; u_{ijct} - \kappa \geq u_{i0ct}\right), \tag{1}$$

[3] The Audit Bureau of Circulations, a nonprofit circulation-auditing organization and my data source for circulation, uses the number of households in a county as the market size for computing county penetration. I follow this and consider a household as a decision unit.

[4] Examples in the literature on multiple discrete choices include Hendel (1999), Nevo, Rubinfeld and McCabe (2005) and Gentzkow (2007). All of these estimate the demand model with individual-level data.

[5] Utility actually varies across i, j, t. The subscript c is redundant in u_{ijct}, as each household can be in only one county. I add the subscript c to emphasize that utility is affected by some county-specific tastes, which are operationalized by county-level demographics.

where the inequality $u_{ij'ct} \geq u_{ijct} \geq \max_{h \neq j'} u_{ihct}$ in the second term ensures that j' is the first best and j is the second best; and κ is a parameter that captures the diminishing utility from subscribing to a second newspaper.

Following the literature, I assume that a household derives utility from some characteristics of a newspaper and that this utility is also affected by county-specific factors and individual-specific tastes. The conditional indirect utility of household i in county c from subscribing to newspaper j in year t is assumed to be

$$u_{ijct} = p_{jt}\alpha_i + \boldsymbol{x}_{jt}\boldsymbol{\beta}_i + \boldsymbol{y}_{jct}\boldsymbol{\psi} + \boldsymbol{z}_{ct}\boldsymbol{\varphi} + \xi_{jct} + \varepsilon_{ijt}, \tag{2}$$

where p_{jt} is the annual subscription price,[6] and $\boldsymbol{x}_{jt}$ is a 3-dimensional vector of the endogenous newspaper characteristics chosen by the newspaper publishers. They are the news hole (non-advertising space), the number of staff for opinion sections, and the number of reporters.[7]

The vector $\boldsymbol{y}_{jct}$ includes the newspaper characteristics that are assumed to be exogenous in the model because they rarely change over time. For example, the location of a newspaper's headquarters determines the distance between the centroid of county c and newspaper j's home county centroid. The distance is included in $\boldsymbol{y}_{jct}$ to capture readers' taste for local newspapers. The vector $\boldsymbol{z}_{ct}$, which includes demographics of county c, captures county/year-specific taste for newspapers. This vector also includes a constant term. Further details concerning the vectors $\boldsymbol{y}_{jct}$ and $\boldsymbol{z}_{ct}$ are provided in Section 4.2 where the estimation results are presented.

The term ξ_{jct} is the unobservable county/year-specific taste for newspaper j. It captures a county-specific taste that is not captured by $\boldsymbol{z}_{ct}$. It also captures characteristics of the newspaper that are relevant for readers but unobservable to the econometrician and therefore not included in $\boldsymbol{x}_{jt}$ or $\boldsymbol{y}_{jct}$; for example, newspaper slant.

It may be interesting to measure slant and study the effect of merger on it. However, I do not have enough data for such a study. A measure of slant based on the similarity of a newspaper's language to that of a congressional Republican or Democrat is proposed by Gentzkow and Shapiro (2010) and relies on catching key phrases in newspapers. The newspapers used in Gentzkow and Shapiro (2010) are newspapers with electronic archives. This paper, however, studies the effect of ownership consolidation and emphasizes the strategic interactions among newspapers. I would have to delete all newspapers without electronic archives and all their competitors from the sample. This deletion would leave very few newspapers in my sample. For this reason, I do not use the measure for slant provided in Gentzkow and Shapiro (2010) and instead let the unobservable term absorb it.

[6] There is very little variation in the newsstand price. For about 80% of daily newspapers, the newsstand price is 50 cents for a daily edition and either \$1.00, \$1.25 or \$1.50 for a Sunday edition. Single-copy sales are around 15% of total circulation.

[7] The news hole consists of opinion sections, stories written by reporters and stories utilized from the external news agencies. Therefore, a newspaper publisher can increase the number of reporters without increasing the news hole by publishing fewer stories from news agencies.

The stochastic term ε_{ijt} is i.i.d. and represents unobservable household specific tastes. Household heterogeneity in price sensitivity and taste for newspaper characteristics is captured by the random coefficients $\alpha_i = \alpha + \sigma_0 \varsigma_{0i}$ and $\boldsymbol{\beta}_i = (\beta_{1i}, \beta_{2i}, \beta_{3i})$, where $\beta_{ki} = \beta_k + \sigma_k \varsigma_{ki}$ is household i's specific taste for the k^{th} endogenous characteristic. I assume that ς_{ki} is identically and independently distributed across characteristics and households, and follows the standard normal distribution. Let $\Phi(\cdot)$ represent the distribution function of $\boldsymbol{\varsigma}_i = (\varsigma_{0i}, \varsigma_{1i}, \varsigma_{2i}, \varsigma_{3i})$.

Instead of treating the utility from the outside choice as fixed, I model it as a time trend to capture changes due to the development of online news sources and the increase in Internet penetration during the sample period. Specifically, I assume that the utility from the outside choice is

$$u_{i0ct} = (t - t_0)\,\rho + \varepsilon_{i0t}, \tag{3}$$

where t_0 is the first year in the data and ρ is a parameter to be estimated.

The market penetration[8] of newspaper j in county c is the aggregation of households' newspaper choices in the county. The aggregation is similar to that in Berry, Levinsohn and Pakes (1995) (henceforth, BLP). Define the "relative" county mean utility, δ_{jct}, as the difference between the mean utility in county c from newspaper j and the mean utility of the outside choice: $\left(p_{jt}\alpha + \boldsymbol{x}_{jt}\boldsymbol{\beta} + \boldsymbol{y}_{jct}\boldsymbol{\psi} + \boldsymbol{z}_{ct}\boldsymbol{\varphi} + \xi_{jct}\right) - (t - t_0)\,\rho$. Then county market penetration can be expressed as a function of $\boldsymbol{\delta}_{ct} = (\delta_{jct}, j = 1, ..., J_{ct})$, $\boldsymbol{p}_{ct} = \left(\boldsymbol{p}_{jt}, j = 1, ..., J_{ct}\right)$ and $\boldsymbol{x}_{ct} = (\boldsymbol{x}_{jt}, j = 1, ..., J_{ct})$: $\mathscr{s}_j(\boldsymbol{\delta}_{ct}, \boldsymbol{p}_{ct}, \boldsymbol{x}_{ct}; \boldsymbol{\sigma}, \kappa)$. See Appendix C for the expression of the county penetration function $\mathscr{s}_j$.

Following BLP, I do not take the market penetration equation to estimation directly, but invert it to obtain the relative mean utility for estimation. The only concern is whether the invertibility result in BLP for a single discrete choice model can be extended to the current multiple discrete choice model. I show in Appendix C that this invertibility holds under two conditions and furthermore, the contraction mapping defined in BLP is still viable, leading to a simple algorithm to solve for $\boldsymbol{\delta}_{ct}$. The two conditions are (i) $0 < s_{jt} < 1$ for $\forall j = 1, ..., J_{ct}$ and (ii) $\sum_{j=1}^{J_{ct}} s_{jt} < 2$. These two assumptions are quite mild. Assumption (i) means that there is always some household choosing newspaper j and some household not choosing it. Assumption (ii) means that there is always some household choosing to purchase fewer than two newspapers. Under these two conditions, the solution to $s_{jct} = \mathscr{s}_j(\boldsymbol{\delta}_{ct}, \boldsymbol{p}_{ct}, \boldsymbol{x}_{ct}; \boldsymbol{\sigma}, \kappa)$ is unique. Denote this solution by $\boldsymbol{\delta}_{ct}(\boldsymbol{s}_{ct}; \boldsymbol{\sigma}, \kappa)$.[9] Therefore, for the true value of the parameters,

$$\delta_{jct}(\boldsymbol{s}_{ct}; \boldsymbol{\sigma}, \kappa) = p_{jt}\alpha + \boldsymbol{x}_{jt}\boldsymbol{\beta} + \boldsymbol{y}_{jct}\boldsymbol{\psi} + \boldsymbol{z}_{ct}\boldsymbol{\varphi} - (t - t_0)\,\rho + \xi_{jct}. \tag{4}$$

This is the first estimation equation. To conclude the description of the demand for newspaper j, let H_{ct} be the number of households in county c in year t. The demand for newspaper j, i.e.,

[8] This is typically called "market share" in a single discrete choice model. But in a multiple discrete choice model, the sum of "market shares" can be larger than 1. "Market penetration" is therefore a better term and is used by the Audit Bureau of Circulations.

[9] The solution depends on $(\boldsymbol{p}_{ct}, \boldsymbol{x}_{ct})$ as well as $\boldsymbol{s}_{ct}$. The subscript ct is added to $\boldsymbol{\delta}$ to recognize this dependence.

the total circulation of newspaper j, is then the sum of the circulation in all counties covered by newspaper j (denoted by $\mathcal{C}_{jt}$):

$$q_j\left(\boldsymbol{\delta}_{ct}, \boldsymbol{p}_{ct}, \boldsymbol{x}_{ct}; \boldsymbol{\sigma}, \kappa\right) = \sum_{c:\ c\in\mathcal{C}_{jt}} H_{ct} \mathscr{s}_j\left(\boldsymbol{\delta}_{ct}, \boldsymbol{p}_{ct}, \boldsymbol{x}_{ct}; \boldsymbol{\sigma}, \kappa\right). \tag{5}$$

The demand for advertising is modeled as in Rysman (2004):

$$a\left(r_{jt}, q_{jt}, \eta_{jt}; \boldsymbol{\lambda}\right) = e^{\eta_{jt}} q_{jt}^{\lambda_1} r_{jt}^{\lambda_2}, \tag{6}$$

where r_{jt} and q_{jt} are the advertising rate in column inches and the total circulation of newspaper j. The demographics of newspaper j's circulation area in year t is captured by η_{jt}. Specifically, I operationalize η_{jt} as follows. Let $\boldsymbol{z}_{ct}\boldsymbol{\phi}$ be a linear combination of observable demographics of county c, where $\boldsymbol{\phi}$ is a vector of parameters to be estimated. Then η_{jt} is defined as the circulation-weighted sum of $\boldsymbol{z}_{ct}\boldsymbol{\phi}$ over the counties covered by newspaper j: $\eta_{jt} = \sum_{c:\ c\in\mathcal{C}_{jt}} \frac{q_{jct}}{q_{jt}} \boldsymbol{z}_{ct}\boldsymbol{\phi}$.

Let ι_{jt} be an i.i.d. and mean zero measurement error for advertising linage; then the second estimation equation is

$$\log a_{jt} = \sum_{c:\ c\in\mathcal{C}_{jt}} \frac{q_{jct}}{q_{jt}} \boldsymbol{z}_{ct}\boldsymbol{\phi} + \lambda_1 \log q_{jt} + \lambda_2 \log r_{jt} + \iota_{jt}. \tag{7}$$

Note that advertising demand depends on circulation. On the other hand, I have assumed that readers only care about the news hole and do not care about advertising. Rysman (2004), in contrast, allows consumers to value advertising in his study of the Yellow Pages market because consumers typically read Yellow Pages for commercial information. In newspaper industry, however, there is mixed evidence on the effect of advertising on circulation.[10] For example, Dertouzos and Trautman (1990) use data for 129 U.S. newspaper firms that responded to a questionnaire in 1980 and find that readers value advertising. Sonnac (2000), on the other hand, presents evidence that most European readers dislike advertisement.

2.2 Supply

The term "market" is typically used to describe either a set of competing firms or a set of available products. This implies that a market is a geographic area that satisfies two criteria: (i) all consumers in the area face the same choice set and (ii) the suppliers of these choices in the area compete with each other and with no one else. In the daily newspaper industry, however, there is no geographic area satisfying both criteria because circulation areas of newspapers partially overlap.

[10] I do not have enough information from my data to estimate the effect of advertising on circulation. Data on advertising quantity is available only for a subset of newspapers. These newspapers are very often in the same counties with newspapers whose advertising linage is not observable. In estimating a discrete choice model allowing consumer heterogeneity, a whole choice set has to be excluded from the estimation when the characteristic of one product is not available. As a result, I do not have enough data to estimate the effect of advertising on circulation.

Table 1 shows the percentage of circulation in the common area (the set of counties where both newspapers are circulated) of two newspapers as a fraction of the total circulation for each member of the pair. It indicates that for 656 *newspaper pairs* in the data for 2005, the overlapping percentage is above 25% for both members. 816 *newspapers* are not in such a pair with any newspaper. When the criterion for overlapping decreases to 20%, the number of overlapping pairs increases to 724.

Table 1: Newspaper Coverage Overlap in 2005

criterion	25%	20%	15%	10%	5%
number of overlapping newspaper pairs	656	724	839	1021	1463
number of newspapers without a significant overlap	816	781	727	631	419

The partial overlapping of newspaper coverage leads to a chain of substitution. For example, when newspapers A and B compete in county 1 and newspapers B and C compete in county 2, the three newspapers are interacting in a single game because A and B, as well as B and C, are direct competitors and A and C are indirect competitors because they share a common competitor.[11] In that sense, all newspapers in the U.S. are potentially competing, which makes the model intractable. To limit the number of players in a game, I make three assumptions.

First, it is unreasonable to think that national newspapers compete with all small newspapers. Thus, I assume that the characteristics and prices of the three national newspapers (Wall Street Journal, New York Times and USA Today) are taken as given in the model.

Second, a newspaper has to report to the Audit Bureau of Circulations, a nonprofit circulation-auditing organization and my data source for circulation, its circulation in all counties receiving 25 or more copies. But a newspaper probably does not compete in a county with such small circulation. Therefore, for each newspaper/year, I sort the counties covered in descending order of county circulation and define the market of the newspaper as the set of counties that covers at least 85 percent of total circulation. I assume that a newspaper only competes with newspapers in this set of counties. This criterion is suggested by the Audit Bureau of Circulations for defining the Newspaper Designated Market. According to the Audit Bureau of Circulations, the Newspaper Designated Market is the "geographical area which is considered to be the market served by the newspaper." For the newspapers whose designated market is directly observable, this criterion is indeed consistent with the market information in the data.

Finally, on the supply side, I assume that a newspaper publisher can exploit economies of scope only if the home counties of its newspapers are in the same Metropolitan Statistical Area (MSA). Details on economies of scope are described later in this section.

[11] Few newspapers have different subscription prices for households outside the home county or the home state. In 2005, for example, less than 30 newspapers among more than 1400 newspapers charge different prices according to the subscriber's location. For these newspapers, I use the local price and ignore price discrimination. Also note that newspaper characteristics studied in this paper are the same no matter where the subscriber of the newspaper is.

These three assumptions limit the number of players in a game. I now describe the game that models the supply side. In the game, all player publishers[12] choose the characteristics of their newspapers in the first stage and newspaper prices and advertising rates in the second stage. In the remainder of this section, the subscript t is suppressed for ease of exposition. It is only restored in the statement of estimation equations.

Suppose the fixed cost of choosing certain combination of newspaper characteristics is given by $fc(\boldsymbol{x}_j, \boldsymbol{\nu}_j; \boldsymbol{\tau})$, where $\boldsymbol{x}_j$ stands for the characteristics, $\boldsymbol{\nu}_j$ represents the unobservable cost shocks and $\boldsymbol{\tau}$ is a vector of parameters. This cost is fixed with respect to circulation and advertising. Then the profit function that is relevant for the first-stage decision is

$$\pi_j^{\mathrm{I}}(\boldsymbol{x}) = \pi_j^{\mathrm{II}}(\boldsymbol{p}^*(\boldsymbol{x}), \boldsymbol{r}^*(\boldsymbol{x}); \boldsymbol{x}) - fc(\boldsymbol{x}_j, \boldsymbol{\nu}_j; \boldsymbol{\tau}), \tag{8}$$

where $\pi_j^{\mathrm{II}}(\boldsymbol{p}, \boldsymbol{r}; \boldsymbol{x})$ is the variable profit from circulation and advertising, and $p_j^*(\boldsymbol{x})$ and $r_j^*(\boldsymbol{x})$ are equilibrium newspaper prices and advertising rates. (In fact, the equilibrium prices also depend on other variables such as the county demographics. They are omitted for presentational simplicity.) The variable profit π_j^{II} is the sum of circulation profit, display advertising profit and preprint profit. I now specify each of the three components.

Circulation profit is the difference between circulation revenue determined by the demand for newspapers described in Section 2.1 and the variable cost of printing and delivery. This cost varies with circulation and at the margin depends on publication frequency and the number of pages. To capture potential economies of scale and economies of scope in printing and delivery, I allow average cost to depend on the total circulation of all newspapers that circulate in the neighboring area of newspaper j (defined by whether their home counties are in the same MSA) and are owned by j's publisher. This total circulation is denoted by Q_j. Specifically, $Q_j = q_j$ when j's publisher owns only one newspaper (i.e., newspaper j) or when the home counties of its other newspapers are not in the same MSA as that of j. Otherwise, Q_j is the total circulation of all of its newspapers whose home counties are in the same MSA as that of newspaper j. To summarize, I assume that the average cost is

$$ac_j^{(q)} = (\gamma_1 + \gamma_2 f_j + \gamma_3 (x_{1j} + a_j)) log(Q_j)^{\gamma_4} + \omega_j, \tag{9}$$

where f_j is the publication frequency measured by the number of issues per year, $(x_{1j} + a_j)$ is the annual pages, i.e., the sum of annual non-advertising space in pages and annual display advertising linage,[13] and ω_j is an unobservable factor that determines the average cost. Note that no other

[12]Under the assumptions above, there might be newspapers that circulate in the market of the players but are not competing with them. They are called "non-players" in this game. For example, the three national newspapers are non-players. Since non-players in a game are assumed not to compete with the player newspapers, their quality and prices are taken as given in the game.

[13]There is a slight abuse of notation here. The annual display advertising linage, a_j, is measured by column inches. According to the *Editor and Publisher International Year Book*, a typical U.S. daily newspaper page has 6 columns with 21-inch depth. Therefore, in fact, the annual pages is $x_{1j} + \frac{a_j}{126}$.

newspaper characteristic besides news hole (x_{1j}) and frequency (f_j) affect the variable cost. That is because the cost of increasing some characteristics of a newspaper, such as the number of reporters, is independent of circulation.

The advertising demand described in Section 2.1 is really demand for display advertising, which is printed on the newspapers' pages along with the news. There exists another type of advertisement, namely preprints, which are inserted into copies of a newspaper and distributed along with them. This is essentially a delivery service provided by newspapers. I do not observe the advertising rate for preprint. Therefore, I do not derive the preprint profit from a demand model. Instead, I assume that it is a simple quadratic function of circulation:

$$\mu_1 q_j + \frac{1}{2}\mu_2 q_j^2. \tag{10}$$

The preprint profit can be considered as a subsidy to production costs. Therefore, the identification of μ_1 and μ_2 is similar to that of the parameters in the variable cost.

Display advertising profit, on the other hand, is derived from the demand model. Display advertising involves two costs. One is the cost of printing, which is captured by the cost varying with circulation as explained above. The other cost is the marginal advertising sales cost, which I assume is

$$mc_j^{(a)} = (1 + 1/\lambda_2)\left(\bar{\zeta} + \zeta_j\right), \tag{11}$$

where λ_2 is the price elasticity of demand for display advertising, as defined in (6),[14] and ζ_j is a mean-zero exogenous random variable. Then the display advertising profit is given by

$$r_j a_j - mc_j^{(a)} a_j. \tag{12}$$

Note that the main arguments offered for ownership consolidation revolve around the concepts of synergies in printing and in the delivery of newspapers. There is rarely a similar discussion on synergies for the advertising sales component. I therefore assume a constant marginal advertising sales cost while allowing the average cost of circulation to vary with circulation.

In summary, the variable profit is given by the sum of circulation profit, display advertising profit and preprint profit:

$$\pi_j^{\mathrm{II}} = \left(p_j q_j - ac_j^{(q)} q_j\right) + \left(r_j a_j - mc_j^{(a)} a_j\right) + \left(\mu_1 q_j + \frac{1}{2}\mu_2 q_j^2\right). \tag{13}$$

Finally, I assume the slope of the fixed cost $fc\left(\boldsymbol{x}_j, \boldsymbol{\nu}_j; \boldsymbol{\tau}\right)$ with respect to the k^{th} endogenous characteristic x_{kj} is

$$\tau_{k0} + \tau_{k1} x_{kj} + \nu_{kj}. \tag{14}$$

The demand for newspapers and display advertising described in Section 2.1 are both annual, i.e., they describe annual subscribers and annual advertising linage. The costs modeled in this section are therefore also annual costs.

[14] λ_2 is added so that the optimal display advertising rate condition (15) is simple.

I conclude this section with a discussion of a Joint Operation Agreement (JOA). There were around 20 JOA's in the U.S. during the data period (1997 – 2005). Newspapers under a JOA combine business operations while maintaining separate and competitive editorial operations. For business operations, the two newspapers under a JOA either form a third company or one of them acts as the operating partner for the other. Therefore, in the model, I assume that the operating party – either the third party or the operating publisher – chooses newspaper subscription prices and advertising rates for both newspapers in the second stage to maximize the joint profit for given newspaper characteristics. In the first stage, the two publishers choose the characteristics of their respective newspapers separately. Since I do not observe how the profit is split between the two newspapers, I assume that each newspaper publisher gets the profit from its own newspaper.

2.3 Necessary Equilibrium Conditions

I now derive the optimality conditions for prices, advertising rates and quality characteristics.[15] Similar to Rosse (1967), these optimality conditions will be used to identify the cost structure of newspaper production.

A newspaper publisher makes a 2-dimensional pricing decision: it must select the subscription price and the display advertising rate for each newspaper it owns. Taking the derivative of the second-stage profit function π_j^{II} in (13) with respect to advertising rate r_j yields the optimal display advertising rate as a function of circulation:

$$r_{jt} = \bar{\zeta} + \frac{\gamma_3}{1 + 1/\lambda_2} log(Q_{jt})^{\gamma_4} q_{jt} + \zeta_{jt}. \tag{15}$$

To derive the first-order condition with respect to subscription price, define a matrix Δ whose (h, j) element is given by

$$\Delta_{hj} = \begin{cases} -\frac{\partial q_h}{\partial p_j}, & \text{if } h \text{ and } j \text{ have the same publisher;} \\ 0 & \text{otherwise.} \end{cases}$$

Similarly, I define the matrix $\mathbf{\Gamma}$, which captures economies of scale and scope, as

$$\Gamma_{hj} = \begin{cases} \frac{\partial ac_h^{(q)}}{\partial Q_h}, & \text{if } h \text{ and } j \text{ have the same publisher \& their home counties are in one MSA;} \\ 0 & \text{otherwise.} \end{cases}$$

Also, I define Λ as a vector of the effect of circulation on display advertising profit:

$$\Lambda_j = -\frac{1}{\lambda_2} \frac{\partial a_j}{\partial q_j} r_j,$$

where λ_2 is the price elasticity of display advertising demand.

[15] I assume that a pure-strategy Nash equilibrium exists. Finding a set of sufficient conditions for the existence of a Nash equilibrium is beyond the scope of this paper.

Then the first-order condition with respect to subscription price can be expressed in matrices as

$$\boldsymbol{p} = \Delta^{-1}\boldsymbol{q} - [\Lambda + (\mu_1 + \mu_2 \boldsymbol{q})] + \Gamma \boldsymbol{q} + \boldsymbol{a}\boldsymbol{c}^{(q)}. \tag{16}$$

The difference between this first-order condition and a standard first-order condition lies in the second term ($[\Lambda + (\mu_1 + \mu_2 \boldsymbol{q})]$), which captures the effect of circulation on total advertising profit, and the third term ($\Gamma \boldsymbol{q}$), which captures economies of scale and scope.

In the first stage, publishers choose newspaper characteristics. The necessary optimality condition for the k^{th} characteristic is

$$\sum_{h \in \mathcal{J}_{mt}} \left(\frac{\partial \pi^{\mathrm{II}}_{ht}}{\partial x_{kjt}} + \sum_{j' \in \mathcal{J}_{g(jt)}} \frac{\partial \pi^{\mathrm{II}}_{ht}}{\partial p_{j't}} \frac{\partial p^*_{j't}}{\partial x_{kjt}} \right) = \tau_{k0} + \tau_{k1} x_{kjt} + \nu_{kjt}, \tag{17}$$

where $\mathcal{J}_{mt}$ is the set of newspapers of j's publisher m in year t and $\mathcal{J}_{g(jt)}$ represents the set of all player newspapers in the game that jt belongs to. The first term on the left hand side is the direct impact of increasing the k^{th} characteristic of newspaper j on the variable profit of newspaper h owned by the same publisher. A change in x_{kjt} also has an indirect effect on the variable profit of newspaper h through an impact on the equilibrium subscription prices for all newspapers in the game. This indirect effect is captured by the second term.

The partial derivatives $\left(\frac{\partial \pi^{\mathrm{II}}_{ht}}{\partial x_{kjt}}, \frac{\partial \pi^{\mathrm{II}}_{ht}}{\partial p_{j't}} \right)$ in (17) can be computed by taking derivatives of the variable profit function (13). The difficulty is computing the gradient of the equilibrium function, $\frac{\partial p^*_{j't}}{\partial x_{kjt}}$. A straightforward way for computing this gradient is to compute the equilibrium function, which then gives information on the gradient. In fact, it is common in the literature to compute the equilibrium of the whole game, i.e., to solve for the equilibrium product characteristics. Therefore, a typical estimation procedure involves a three-level nested algorithm: in the inner loop, the pricing equilibrium is solved for given product characteristics and model parameters; in the middle loop, the product choice equilibrium is solved for given model parameters; and in the outer loop, parameters are searched to minimize some estimation criterion function. The computational burden of such a nested fixed point problem is nontrivial. As a result, researchers typically use this method to study an industry with a simple market structure, such as the monopoly markets in the cable industry in Crawford and Shum (2006), or an industry where the possible choices for product characteristics are discrete and finite, such as the choice for ice cream flavors in Draganska, Mazzeo and Seim (2009).

I take a different approach. Since the estimation equation (17) is the optimality condition for the *observed* product characteristics, only the values of the gradient at the data points are needed to formulate (17). Therefore, it is not necessary to compute the equilibrium price for each possible quality choice. I compute the values of the gradient at the data points by taking the total derivative of the first-order condition with respect to newspaper prices (16).[16] This approach, however, does require that the profit function has to be differentiable in characteristics. Also, the first-order conditions of prices contain the first-order partial derivatives of the profit function. Total differentiation of these conditions thus involves the second-order partial derivatives of the profit function. This requires that the model captures even the second-order derivative of newspaper publishers' profit structure accurately. Note that the algorithm in the literature described above also requires that the model captures the true profit function accurately so that the equilibrium price function can be accurate. To ease this concern, I collect as much information on the determinants of a newspaper's profit as possible.

3 Data

For this study, I have compiled a new data set from various sources. The data covers all daily newspapers in the United States from 1997 to 2005. Specifically, the data set contains information on quantities and prices on both sides of the market. On the readers' side, I observe county circulation and annual subscription price (q_{jct}, p_{jt}). On the advertisers' side, I observe annual display advertising linage and display advertising rate (a_{jt}, r_{jt}). See Appendix A for a detailed explanation of the data sources and the variable definitions.

The data set also contains information on newspaper characteristics. A newspaper is described by the following attributes: the news hole, the number of opinion section staff, the number of reporters, the frequency of publication, and edition (morning or evening newspaper). Data on these attributes except the news hole is available. The news hole is the non-advertising space in a newspaper. Therefore, the annual news hole can be replaced by $n_j - a\left(r_j, q_j, \eta_j; \boldsymbol{\lambda}\right)$ in the estimation, where n_j is the annual number of pages. The latter depends on observable variables and model parameters.

The number of reporters in this study is measured by the number of reporters weighted by the inverse of the number of titles that each reporter has. For example, if a reporter also holds some managing job, this reporter contributes 1/2 to the number of reporters. The number of opinion section staff is similarly defined.

Data on all variables except display advertising linage, annual subscription price and pages per issue are available for all newspapers during the data period. Display advertising linage data is available for 485 newspaper/years between 1999 and 2005. Information on this subset of newspapers

[16] Abstractly, suppose $F\left(\boldsymbol{x}, \boldsymbol{p}^*\left(\boldsymbol{x}\right)\right) = 0$ is the first-order condition with respect to prices. Then the total derivative gives $\nabla_x F\left(\boldsymbol{x}, \boldsymbol{p}^*\left(\boldsymbol{x}\right)\right) + \nabla_p F\left(\boldsymbol{x}, \boldsymbol{p}^*\left(\boldsymbol{x}\right)\right) \nabla_x \boldsymbol{p}^*\left(\boldsymbol{x}\right) = 0$. Even though the equilibrium function $p_j^*\left(\cdot\right)$ itself is unknown and hard to compute, the value of it at the observed characteristics $\boldsymbol{x}$ is known: it is the observed price p_j. Therefore, $\nabla_x F\left(\boldsymbol{x}, \boldsymbol{p}^*\left(\boldsymbol{x}\right)\right) = \nabla_x F\left(\boldsymbol{x}, \boldsymbol{p}\right)$ and $\nabla_p F\left(\boldsymbol{x}, \boldsymbol{p}^*\left(\boldsymbol{x}\right)\right) = \nabla_p F\left(\boldsymbol{x}, \boldsymbol{p}\right)$ at observed $\left(\boldsymbol{x}, \boldsymbol{p}\right)$. Thus, $\nabla_x \boldsymbol{p}^*\left(\boldsymbol{x}\right)$ at observed $\boldsymbol{x}$ can be easily computed.

is used to identify the advertising demand parameters.[17] Missing data on price or pages per issue (which are newspaper attributes) lead to deletion of observations: all newspapers interacting with a newspaper with missing information on price or pages are deleted from the sample.[18] There are 1387 newspaper/years with missing data on price or pages, leading to the deletion of 6283 newspaper/years, with 6566 newspaper/years remaining. These newspaper/years' markets consist of 9644 county/years. Summary statistics for the main variables for the final sample are provided in Tables 2 and 3.

Table 2: Summary Statistics of Player Newspapers

	mean	median	s.d.	min	max	obs
total circulation	23,075	9,867	44,776	1,132	783,212	6566[a]
price of newspapers ($)	101.52	97.21	33.83	13.19	365.31	
display advertising rate ($/column inch)	27.54	13.44	47.56	2.45	748.70	
frequency (issues/52 weeks)	310.46	312	53.94	208	364	
pages (pages/issue)	29.09	24	20.98	8	254.57	
opinion staff	1.66	0.5	3.83	0	57.53	
reporters	6.12	2	13.93	0	315.63	
market penetration (%)	19.03	11.67	18.57	0.30	97.08	24908[b]
distance to home county (100km)	0.71	0.47	0.81	0	6.64	

[a]These observations are at the newspaper/year level.
[b]These observations are at the newspaper/county/year level.

Table 3: Summary Statistics: County Demographics

	mean	median	s.d.	min	max	obs
% of pop over 25 with bachelor's degree or higher	17.15	15.23	7.28	5.64	60.48	9664[a]
median income ($1,000)	34.32	32.92	7.34	16.36	80.12	
median age	36.53	36.70	3.79	20.70	54.30	
urbanization (%)	50.01	51.23	26.62	0	100	
number of households	36,045	15,326	84,368	710	3,282,266	

[a]These observations are at the county/year level.

4 Estimation

Five estimation equations are taken to the data: (4), (7), (15), (16) and (17). They are respectively derived from newspaper demand, advertising demand and the first-order conditions with respect to advertising rate, subscription price and newspaper characteristics.

[17]Note that I assume that newspapers do not have direct strategic interaction on the advertising side.

[18]This is because, for example, when the price of newspaper j is not observable, the optimality condition for any newspaper j' in j's game is not well-defined. Therefore, j's game is deleted.

The parameters to be estimated include (i) the parameters in the newspaper demand function; (ii) the parameters in the display advertising demand function; (iii) the cost parameters; and (iv) the parameters in the preprint profit function.

The identification of newspaper demand parameters is similar to the identification of analogous parameters in BLP. However, unlike BLP, product characteristics are endogenous in this paper. As explained in Section 4.1, I therefore use a different exogenous variation to identify the effects of product characteristics and prices.

Among other parameters, the identification of the diminishing utility parameter κ needs an explanation. Identification of κ comes from variation in the number of newspapers in a county. In counties with only one newspaper, diminishing utility does not play a role in determining market penetrations. Suppose all parameters were identified using the data from such counties only. Then, based on these estimates, market penetrations in counties with multiple newspapers could be computed assuming that each household chooses *at most one newspaper*. The difference between the observed data and these counterfactual market penetrations assuming a single choice is then explained by the choice of a second newspaper, the probability of which is influenced by κ.

4.1 Instruments

In the model, newspaper publishers know the unobservable (to econometricians) newspaper-county specific taste ξ_{jct} and the unobservable cost shocks $\left(\zeta_{jt}, \omega_{jt}, \boldsymbol{\nu}_{jt}\right)$ before they choose the characteristics, the subscription prices and the advertising rates of their newspapers. These choices are therefore likely to be correlated with the unobservables. Instrumental variables are used to deal with this endogeneity. Specifically, I use the demographics in the market of j's competitors (excluding j's own market) as instruments.

The intuition for why the demographics in competitors' market can be used as instruments is illustrated in Figure 1. The demographics in county 2 influence the demand for newspaper B as well

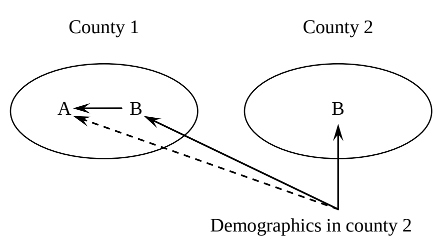


Figure 1: Instruments

as its advertising demand, and thus affect the prices and the attributes of this newspaper. Because newspapers A and B are competitors, B's decision on product characteristics and prices affects A's decision. Therefore, the demographics in county 2 indirectly affect newspaper A's product choice and price decisions. For example, a local newspaper in a small county close to a large city with a metropolitan newspaper might want to position itself as an inexpensive and low-quality newspaper.

This choice of instrument is in the same vein as that in BLP, who uses the characteristics of competitors' products as instruments. The instruments used in BLP are valid because firms consider what kind of products are available in the market when making a price decision and the product characteristics are assumed to be exogenous in BLP. In this paper, the product characteristics are the focus and considered endogenous. But firms consider what kind of consumers they serve, i.e., demographics, when making a decision on product characteristics and prices. The demographics of competitors' markets therefore can be used as instruments. The underlying assumption in BLP is that the product characteristics are exogenous. The underlying assumption here is that the entry/location choices are exogenous. This is plausible because location decisions are typically of a longer horizon than both quality and price decisions.

In summary, the partial overlapping feature of the industry allows the demographics in competitors' markets to be used as instruments, specifically, the excluded instruments. The included instruments include the demographics of a newspaper's own market. Table 4 reports the correlation between the included and the excluded instruments. Specifically, it reports the correlation between the mean educational level, for example, in the market of a newspaper and the mean of the educational levels in the counties that belong to its competitors' markets but are not in its own market. This table shows that the demographics of neighboring counties are not highly correlated, i.e., the included instruments and the excluded instruments are not highly correlated.

Table 4: Correlation of Demographics in Neighboring Counties

	educational level	median income	median age	urbanization
correlation	0.1659	0.3455	0.4652	0.3337

Note that among the demographic measures only the number of households in a county varies across years. This is because the data on the number of households comes from the yearly County Penetration Report by the Audit Bureau of Circulations while the county-level demographics data comes from Census and yearly data is not available. So the main variation is cross-sectional. The exogenous sources of variation that lead to changes in prices and newspaper characteristics over time include the variation in market structure such as ownership and the time trend.

4.2 Estimation Results

Estimation results are presented in Table 5. The endogenous newspaper characteristics include news hole, the number of staff for the opinion sections and the number of reporters.[19] The estimates

[19] In the estimation, I replace x_{kjt} in the utility function by $\log(1 + x_{kjt})$, as this specification of newspaper characteristics explains the data better. In the cost function, I use x_{kjt}.

Table 5: Estimation Results

		parameter	estimate	s.e.
Utility	price ($100)	α	-0.884**	0.014
	log(1+newshole), mean	β_1	0.534**	0.088
	log(1+opinion), mean	β_2	0.487**	0.071
	log(1+reporter), mean	β_3	0.086**	0.043
	price, s.d.	σ_0	0.002**	0.0003
	log(1+newshole), s.d.	σ_1	0.069**	0.017
	log(1+opinion), s.d.	σ_2	0.200**	0.102
	log(1+reporter), s.d.	σ_3	0.019	0.295
	log(households in the market)	ψ_1	-66.132**	5.995
	morning edition	ψ_2	0.468**	0.060
	local dummy	ψ_3	1.811**	0.037
	county distance (1000km)	ψ_4	-4.517**	0.441
	constant	φ_0	-0.090	0.541
	education	φ_1	1.146**	0.210
	median income ($10000)	φ_2	-0.578**	0.204
	median age	φ_3	2.909**	0.602
	urbanization	φ_4	-0.165*	0.095
	time	ρ	0.074**	0.009
	diminishing utility	κ	3.783**	0.780
Display ad demand	total circulation	λ_1	1.870**	0.002
	ad rate	λ_2	-1.154**	0.028
	constant	ϕ_1	-2.695**	0.119
	median income ($10000)	ϕ_2	0.127	0.180
Avg cost of circulation	constant	γ_1	-24.007	1908
	frequency	γ_2	14.884**	7.500
	pages in a year	γ_3	0.058	0.042
	economies of scale/scope	γ_4	-0.987**	0.235
Marginal cost of ad sales		$\bar{\zeta}$	14.423**	0.902
Slope of the fixed cost for opinion staff	constant	τ_{20}	306100**	60293
	opinion	τ_{21}	102450**	38824
Slope of the fixed cost for reporters	constant	τ_{30}	42123*	22505
	reporter	τ_{31}	775	645
Preprint profit	circulation	μ_1	527.830**	235.270
	square of circulation	μ_2	-0.010**	0.001

** indicates 95% level of significance.

* indicates 90% level of significance.

of mean taste parameters ($\boldsymbol{\beta}$) and disutility from price (α) imply that doubling the news hole of a newspaper while also increasing its annual subscription price by 42 dollars leaves mean utility unchanged. Since the estimated reader heterogeneity ($\boldsymbol{\sigma}$) is small, this also means the demand for newspapers would not change much in such a scenario. Similarly, decreasing the number of opinion section staff by half is equivalent to increasing the subscription price by 38 dollars, and decreasing the number of reporters by half – while keeping all else equal, including the news hole size – is tantamount to increasing price by 7 dollars.

The exogenous characteristics include the number of households in the market of a newspaper, whether the newspaper is a morning newspaper or an evening newspaper, whether a county is its home county and the distance between the county and its home county. The corresponding parameters are ψ_1 to ψ_4. The negative sign of ψ_1 indicates that readers value a newspaper with, for example, 10 reporters covering a small region more than they do a newspaper that has 10 reporters and serves a large area. The estimates also show that readers prefer morning newspapers (see the estimate of ψ_2). Readers' taste for local newspapers is captured by the distance between the centroid of county c and the centroid of newspaper j's home county. The local dummy, i.e., whether the distance is 0, is also included to allow readers' taste to be nonlinear in the distance. The estimates of ψ_3 and ψ_4 indicate that readers value newspapers whose home counties are close and have a particular taste for local newspapers.

County demographics used in this paper include educational level, median income, median age and urbanization, of which educational level and age positively affect the demand for newspapers. The positive sign of ρ indicates that readers' utility from subscribing to a newspaper decreases over time. This is consistent with the advent of online news, which motivates the inclusion of the time trend in the model.

The parameter κ measures the diminished utility of subscribing to a second newspaper. In a single discrete choice model, this parameter is essentially set to infinity so that consumers buy at most one product. The estimate of κ in the present multiple discrete choice model is 3.783. It implies that in most county/years, duplicate readership is negligible. In the 476 county/years with a nontrivial percentage of households who choose to purchase two newspapers, on average only 5% of the subscribers do so.

All parameters in the advertising demand function have the expected signs: an increase in circulation and a decrease in advertising rate raise advertising demand. The price elasticity of display advertising demand is close to -1. The elasticity with respect to circulation, however, is larger than 1. As will be explained in the next section, this has an important implication for how publishers adjust the characteristics and price of their newspapers after a market structure change.

I set the parameters in the cost of increasing the news hole (τ_{10}, τ_{11}) to zero because specifications that do not restrict these parameters indicate that the news hole does not affect the fixed costs (fixed with respect to circulation). This can be explained as follows. The news hole consists of stories written by reporters and those bought from news agencies. The former can be increased by hiring more reporters. But this effect on fixed costs is already captured by the cost of reporters. The cost of the latter is de facto a variable cost, because news agencies typically set their rates based on the circulation of a newspaper instead of the number of stories that the newspaper buys.

Synergies in printing and delivery are very often the argument for an ownership consolidation. The negative sign of the estimate of γ_4 is consistent with this argument.

5 Case Studies in the Minneapolis/St. Paul Metropolitan Area

In this section, I study a merger of two direct competitors in the Minneapolis market that was blocked by the Department of Justice. I also analyze the effect of a merger of two newspapers in this market that do not compete directly, but share a common competitor. The effects of ownership consolidation on the product choice and price decisions of newspaper publishers as well as the resulting welfare implications are investigated.

A brief discussion of welfare measures is in order (details can be found in Appendix B). The welfare effect on readers is measured by compensating variation similar to that in Small and Rosen (1981). Publisher surplus is given by the profit function in (8). Advertiser surplus, however, cannot be estimated. Since I observe only the advertising linage for newspapers instead of each advertiser's individual behavior, only the price elasticity of the market demand for advertising is identified. Due to the potential externality of aggregate advertising on the effectiveness of individual advertising, the market demand does not correspond to an individual agent's willingness to pay. Thus there is not enough information in the data to measure advertiser surplus. However, as shown in Appendix B, there is enough information to compute the percentage change of advertiser surplus after an ownership consolidation.

Throughout this section, I use "ownership consolidation" and "merger" interchangeably.

Case 1. Ownership Consolidation of Direct Competitors

In 2006, the McClatchy Company purchased its much larger rival Knight Ridder Inc. After the acquisition of Knight Ridder, McClatchy owned two daily newspapers in the Minneapolis/St. Paul metropolitan area: the *Minneapolis Star Tribune* (henceforth, *Star*) and the *St. Paul Pioneer Press* (henceforth, *Pioneer*), the latter of which was previously owned by Knight Ridder. Three months after the announcement of the transaction, the Department of Justice filed a complaint. Two months later, McClatchy sold *Pioneer* to the Hearst Corporation, which later sold it to MediaNews Group. Neither Knight Ridder nor MediaNews owned another newspaper in this market. Therefore, this series of transactions did not lead to a market structure change in the framework of this paper, as the publisher of *Pioneer* was simply relabeled.

In this section, I investigate what would have happened to newspaper quality, subscription prices, advertising rates and welfare if the ownership consolidation of *Star* and *Pioneer* had been upheld. These two newspapers are in a game with two other newspapers: the *Faribault Daily News* and the *St. Cloud Times*. Their markets are illustrated in Figure 2. The Minneapolis-based *Star* and the St. Paul-based *Pioneer* are direct competitors as their markets overlap in five counties. *Star* circulates in a larger area. The *Faribault Daily News* and the *St. Cloud Times* compete with *Star* only.

Tables 6 and 7 present newspaper quality characteristics, subscription prices and advertising rates at the post-merger equilibrium[20] when only prices are adjusted (Table 6) and when both quality and prices are endogenously chosen by publishers (Table 7). Standard deviations of the change in prices and characteristics are also reported.

Table 6: Effects of Ownership Consolidation of *Star Tribune* and *Pioneer Press* without Quality Adjustment

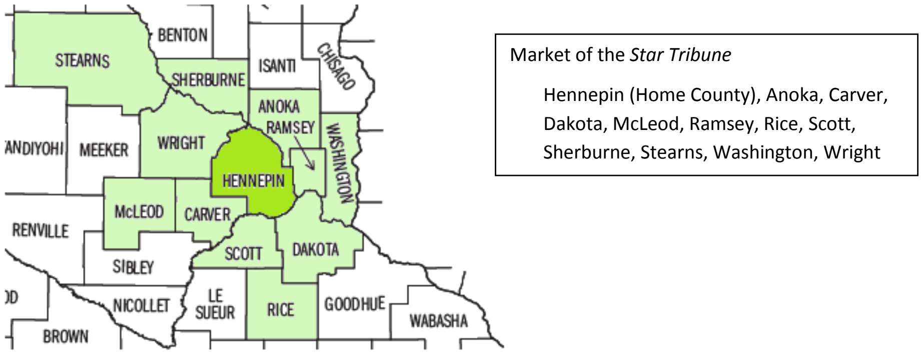
	price ($/year)				ad rate ($/column inch)				circulation			
	before	after	change	s.d.	before	after	change	s.d.	before	after	change	s.d.
Star Tribune	173	180	7	0.81	230.88	226.45	-4.43	0.89	317337	310418	-6919	585
Pioneer Press	172	188	16	4.13	153.08	146.37	-6.71	1.59	159864	149344	-10520	3172
Faribault Daily	111	111	0	0.09	12.37	12.40	0.03	0.01	6384	6480	96	14
St. Cloud Times	150	150	0	0.07	44.15	44.19	0.04	0.01	24578	24696	118	18

Table 7: Effects of Ownership Consolidation of *Star Tribune* and *Pioneer Press* with Quality Adjustment

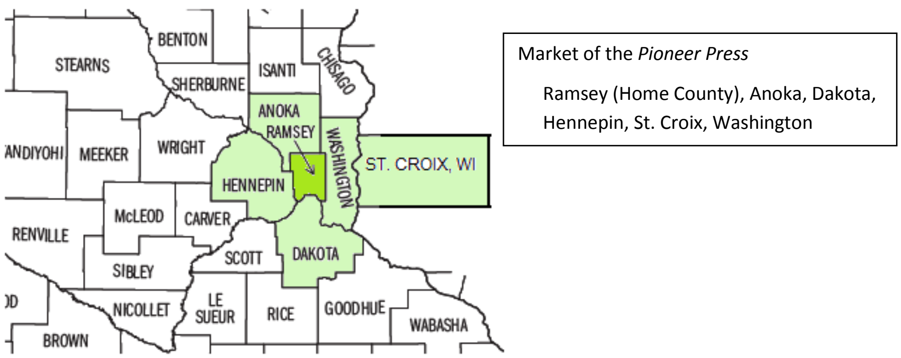
	news hole (pages/year)				opinion				reporter			
	before	after	change	s.d.	before	after	change	s.d.	before	after	change	s.d.
Star Tribune	14300	19303	5003	2100	29.08	30.06	0.98	0.38	115.92	121.65	5.73	6.42
Pioneer Press	14195	14191	-4	605	19.92	18.90	-1.02	0.19	68.58	64.53	-4.05	1.24
Faribault Daily	7258	7128	-130	62	0	0	0	0.01	1	1	0	0.02
St. Cloud Times	15009	15348	339	229	2.33	2.33	0	0.02	10	9.97	-0.03	0.11
	price ($/year)				ad rate ($/column inch)				circulation			
Star Tribune	173	205	32	6.79	230.88	226.63	-4.25	1.12	317337	310997	-6340	2181
Pioneer Press	172	185	13	3.99	153.08	146.03	-7.05	1.13	159864	148096	-11768	2150
Faribault Daily	111	110	-1	0.56	12.37	12.40	0.03	0.02	6384	6466	82	54
St. Cloud Times	150	152	2	2.03	44.15	44.17	0.02	0.06	24578	24619	41	163

From Table 6, we can see that when one assumes no quality adjustment, both *Star* and *Pioneer* increase their subscription prices. This is because after the publisher of *Star*, McClatchy, purchases *Pioneer*, it internalizes the positive cross price effect of these two newspapers: a higher price for *Star*, for example, leads to an increase in the market share of *Pioneer* and therefore raises its profit.

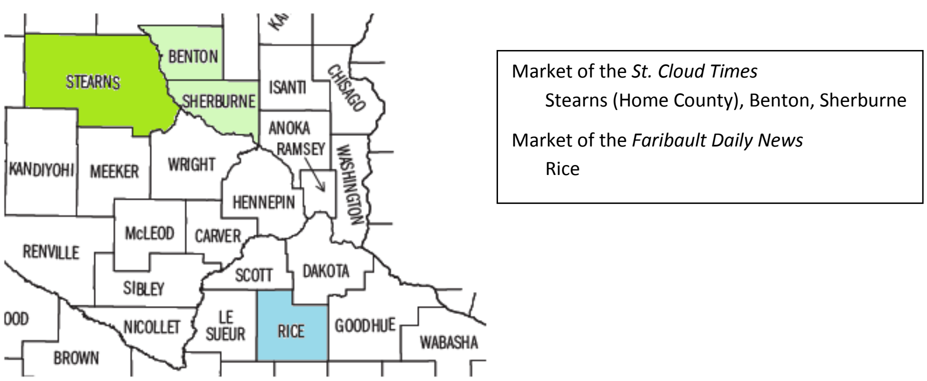
[20] Proving the uniqueness of an equilibrium is beyond the scope of this paper. I "verify" the uniqueness by using different algorithms and trying different starting points for computing the equilibrium.



(a) Market of the *Star Tribune*



(b) Market of the *Pioneer Press*



(c) Markets of the *St. Cloud Times* and the *Faribault Daily News*

Figure 2: Newspaper Designated Market

The table also shows that the price of *Pioneer* is increased by 16 dollars, larger than the price adjustment for *Star*. In other words, the adjustment of the smaller party to the merger (*Pioneer*) is much larger than that of the larger party (*Star*). As explained in Section 4.2, the model estimates indicate that the advertising profit function is convex in circulation, implying that the marginal value of circulation is higher for larger newspapers. Therefore, a multi-newspaper publisher has an incentive to shift circulation from its small to its large newspapers. Here, McClatchy who owns both *Star* and *Pioneer* after the ownership consolidation adjusts the price of the smaller newspaper by a bigger margin due to this incentive.

Table 7 shows what happens when product characteristics adjustments are allowed. When a multiple-product firm chooses the position of its products, on the one hand, it wants to keep its products apart due to concerns of cannibalization, but on the other hand, it does not want to leave space for competitors to move their products in between. In the market studied in this section, competition concern is insignificant because the overlap between the two newspapers in the merger and the other newspapers is small. Hence, the cannibalization concern dominates. Therefore, McClatchy increases the product differentiation between *Star* and *Pioneer*, which is consistent with the simulation result in Table 7 that the characteristics of *Star* increase, while those of *Pioneer* decline. The increased quality of *Star* allows McClatchy to increase its price more than it would if there were no adjustment in characteristics: the price of *Star* is raised by 32 dollars in Table 7 as opposed to seven dollars in Table 6. So ignoring quality adjustment leads to an underestimation of the price change of one newspaper by 25 dollars, which is 25% of the average annual subscription price. (The average annual subscription price is 101 dollars.) Table 7 also shows that the overall change in characteristics and prices leads to a decline in circulation of both newspapers, with a larger drop in *Pioneer* due to the aforementioned incentive to shift circulation from the smaller newspaper to the larger newspaper.

Overall, reader surplus declines by 4.02 million dollars and publisher surplus increases by 15.03 million dollars. The large increase in publisher surplus is due to economies of scope after the merger of *Star* and *Pioneer* as well as the change in market power. As explained at the beginning of Section 5, there is only enough information to measure the percentage change of advertiser surplus. In this study, advertiser surplus declines by 4.49%. The welfare change *without* quality adjustment is -4.04 million dollars for readers, -5.43% for advertisers and 13.96 million dollars for publishers. Therefore, in this market, the combined change of quality and price leaves reader surplus almost the same as in the scenario where only prices are adjusted. However, ignoring quality adjustment overestimates the loss in advertiser surplus and underestimates the gain in publisher surplus in this particular market. The general relationship between the bias in estimating the welfare effect from ignoring quality adjustment and the underlying market structure is analyzed in Section 6.

Case 2. Ownership Consolidation of Indirect Competitors

In the above ownership consolidation study, the two parties to the merger are direct competitors. This is usually the main focus in policy analyses. In fact, similar quality and price cross-effects exist even when the merged parties just share a common competitor. To illustrate this point and quantify the effect, I study a counterfactual merger of the *Pioneer Press* and the *St. Cloud Times*, whose markets do not overlap as Figures 2(b) and 2(c) show.

The results are presented in Table 8. Even though *Pioneer* and the *St. Cloud Times* do not overlap in their markets, they both compete with *Star*. As a result, a change in the price or characteristics of *Pioneer* can lead to a change in the best response of *Star* and hence affects the profit of the *St. Cloud Times*. Therefore, after the ownership consolidation of *Pioneer* and the *St. Cloud Times*, the publisher adjusts the characteristics and prices of these two newspapers to internalize the cross effects. Table 8 confirms this adjustment, even though the adjustment of the *St. Cloud Times* is statistically insignificant. The *Star Tribune*, who competes with both newspapers, also reacts. In fact, a comparison of Table 7 (the effect of ownership consolidation of *Star* and *Pioneer*) and Table 8 (the effect of ownership consolidation of *Pioneer* and the *St. Cloud Times*) shows that the quality adjustment of *Star* is comparable in these two counterfactual consolidations. The overall welfare impact of such an ownership consolidation is, however, much smaller than that of merging two direct competitors. Reader surplus decreases by 8542 dollars, publisher surplus increases by 0.94 million dollars and the change in advertiser surplus is negligible.

Table 8: The Effect of the Ownership Consolidation of *Pioneer Press* and *St. Cloud Times*

	news hole (pages/year)				opinion				reporter			
	before	after	change	s.d.	before	after	change	s.d.	before	after	change	s.d.
Star Tribune	14300	18693	4393	1191	29.08	30.25	1.17	0.34	115.92	122.22	6.30	4.24
Pioneer Press	14195	13363	-832	276	19.92	19.26	-0.66	0.21	68.58	65.71	-2.87	2.02
Faribault Daily	7258	7078	-180	21	0	0	0	0.01	1	0.97	-0.03	0.02
St. Cloud Times	15009	15427	418	3317	2.33	2.25	-0.09	0.1	10	9.67	-0.33	0.62
	price ($/year)				ad rate ($/column inch)				circulation			
Star Tribune	173	196	23	3.39	230.88	231.1	0.22	0.49	317337	318200	863	1976
Pioneer Press	172	166	-6	1.01	153.08	152.73	-0.35	0.16	159864	158580	-1284	713
Faribault Daily	111	110	-1	0.40	12.37	12.36	-0.01	0.01	6384	6352	-32	40
St. Cloud Times	150	153	3	20.06	44.15	44.04	-0.11	0.23	24578	24216	-362	758

6 Welfare Analysis of Duopoly Mergers and Triopoly Mergers

Welfare effects of ownership consolidation in a market depend on the details of the market structure. The last section shows how the framework in this paper can be used to study ownership consolidation in one specific market.[21] I now investigate the general pattern of how the welfare effect of a merger varies with market characteristics. The results of this investigation can be used to guide policy. To this end, I study the welfare implications of ownership consolidations in duopoly and triopoly markets. I will present the distribution of the welfare effects for such mergers in all duopoly and triopoly markets in the 2005 sample, and then examine how they vary with market characteristics. In a duopoly merger, the publisher of one newspaper buys the other and becomes a monopolist in the market. A triopoly merger is defined as one in which the two largest newspapers merge.

Figures 3 and 4 show welfare changes after an ownership consolidation in 39 duopoly markets and 13 triopoly markets in the 2005 sample, the last year in the data. The markets are sorted according to the change in average per-household reader surplus ($\overline{\Delta RS}$) given quality adjustment. Dots in Figure 3(a) represent $\overline{\Delta RS}$ in simulations where quality adjustments are allowed. Asterisks represent $\overline{\Delta RS}$ when such quality adjustments are not allowed. The distance between an asterisk and a dot on the same vertical line therefore represents the bias in estimating $\overline{\Delta RS}$ when quality adjustment is ignored. Figure 3(b) plots changes in total reader surplus (ΔRS). Figures 3(c) and 3(d) show percentage changes in advertiser surplus (ΔAS) and changes in publisher surplus (ΔPS) in millions, respectively. Finally, Figure 4 represents the same measures for the 13 triopoly markets. The mean changes in different welfare measures are presented in Table 9. Reader surplus falls in 35 duopoly markets and 12 triopoly markets. Synergy allows an increase in reader surplus in a few markets after ownership consolidation. Producer surplus increases in all duopoly markets as expected, and decreases in 2 out of 13 triopoly markets due to the existence of a competitor.

Table 9: Mean Welfare Changes across Duopoly and Triopoly Markets

	$\overline{\Delta RS}$ \$	ΔRS \$(millions)	$\%\Delta AS$ %	ΔPS \$(millions)
Duopoly with quality adjustment	-12.31	-0.65	-0.35	0.81
Duopoly without quality adjustment	-11.81	-0.57	-9.86	0.43
Triopoly with quality adjustment	-6.42	-0.33	-2.08	1.16
Triopoly without quality adjustment	-5.46	-0.18	-7.38	0.77

Two variations of the welfare effects across markets as shown in Figures 3 and 4 are studied: (i) variation of the welfare change with quality adjustment and (ii) variation of the magnitude of the bias. I focus on reader surplus here. Let $\overline{\Delta RS}_m$ be the change in average per-household reader

[21] The performance of the framework can be tested by comparing the simulation result of ownership consolidation that happened during the data period with the observable outcomes of the mergers. But note that the simulation is based on the assumption that the unobservable cost shocks and taste shocks do not change. So any discrepancy between simulation results and data could be explained by perturbations in these unobservables. I therefore need a decent amount of real consolidations to test the performance of the framework. Unfortunately, the sample includes too few consolidations to make this type of assessment feasible here.

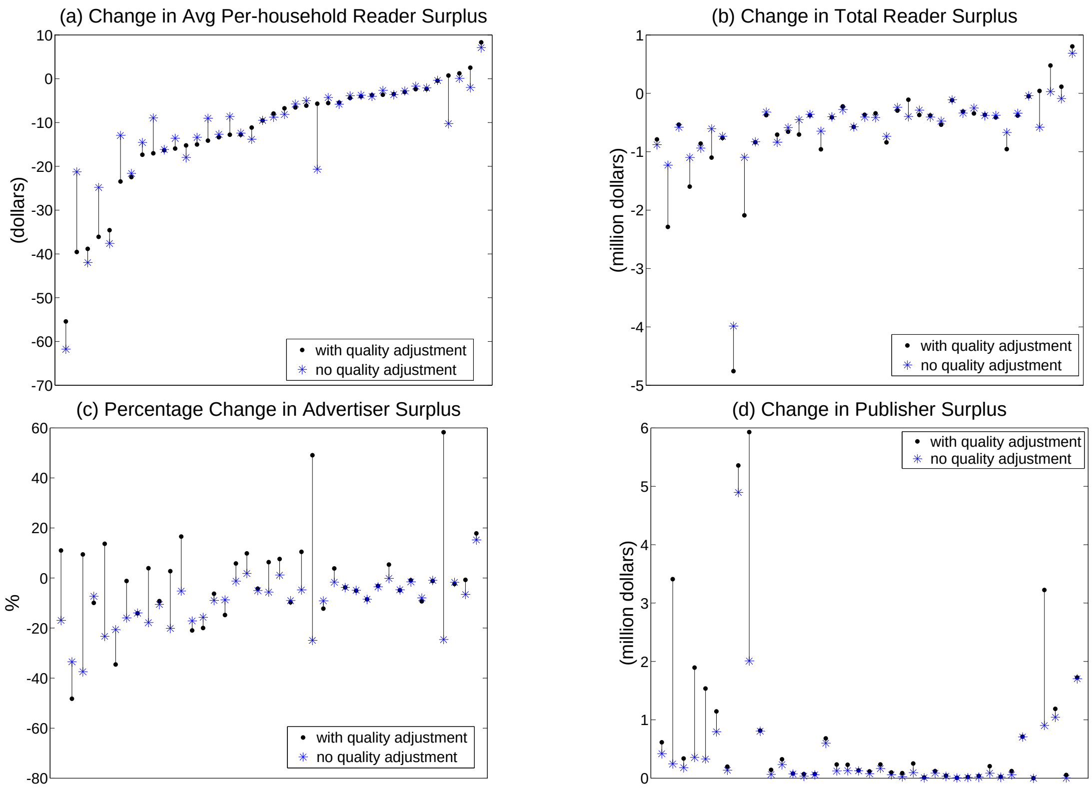


Figure 3: Welfare Implications of Duopoly Mergers

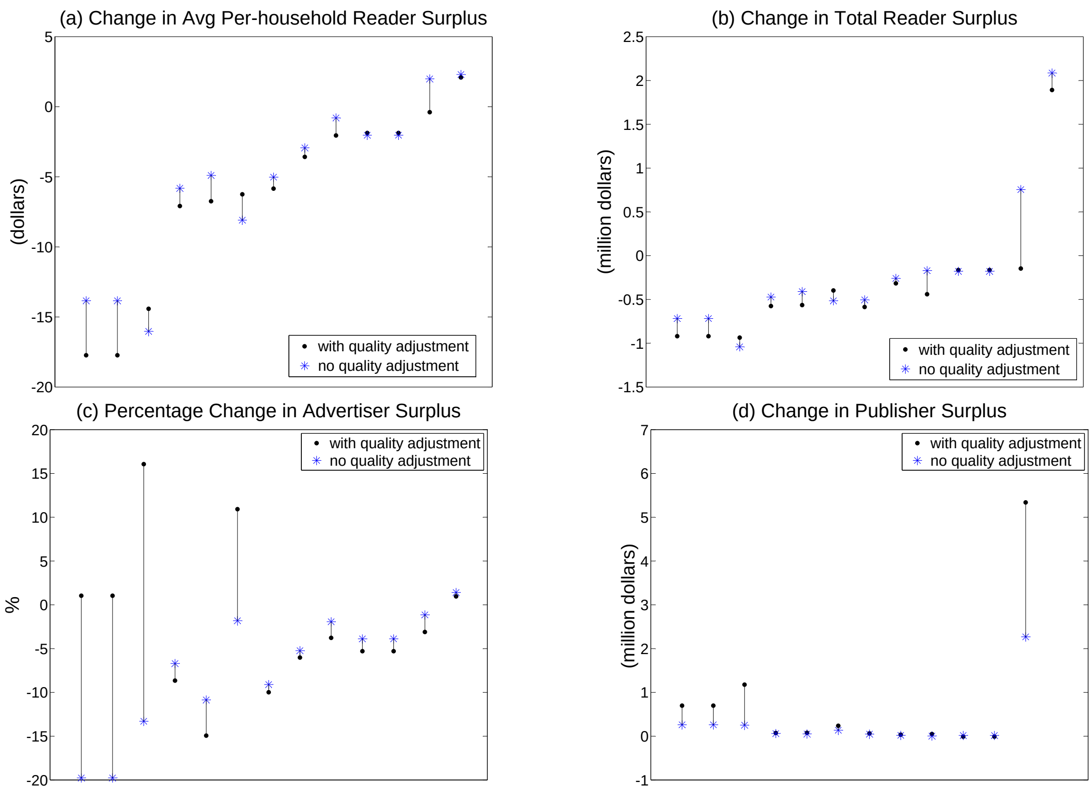


Figure 4: Welfare Implications of Triopoly Mergers

surplus in market m, and $\Delta\overline{\Delta RS}_m$ be the distance between $\left(\overline{\Delta RS}_m \text{ without quality adjustment}\right)$ and $\left(\overline{\Delta RS}_m \text{ with quality adjustment}\right)$. To understand the above variations, I run two regressions of $\overline{\Delta RS}_m$ and $\Delta\overline{\Delta RS}_m$ on some market characteristic variables. A market in the newspaper industry is characterized by market size, demographics of readers, the set of newspapers, each newspaper's designated market, ownership structure, etc. Since the market structure cannot be captured by simple indices, I regress $\overline{\Delta RS}_m$ and $\Delta\overline{\Delta RS}_m$ on a triopoly dummy and endogenous variables that are correlated with the underlying market characteristics. The regression therefore shows a correlation pattern rather than a causal effect.

The results of the first regression are as follows (standard errors are in parentheses):

$$\begin{array}{llllll} \overline{\Delta RS}_m = 18.15 & -\ 68.79\,\text{pen}_m & -\ 8.73\,\text{overlap}_m & +\ 3.81\log\left(\frac{q_{1m}}{q_{2m}}\right) & +\ 4.85\,\text{triopoly}_m + \varrho_m. \\ \qquad\quad (6.04) & \quad (11.92) & \quad (3.32) & \qquad (1.53) & \qquad (2.43) \end{array}$$

The impact of ownership consolidation on readers' welfare depends on how much readers in a market value newspapers in general. Obviously, if households in a market do not like reading newspapers, then changes in newspaper quality and prices do not affect their welfare much. The pre-merger newspaper penetration (pen_m), measured by the ratio of the total newspaper circulation to the number of households in market m, is used to capture this aspect of the market. The negative sign in the estimate is as expected: readers' welfare loss ($-\overline{\Delta RS}_m$) increases when readers care more about newspapers. An increase in the penetration by 1 percentage point is related to an increase in the average welfare loss per household of 69 cents.

Another market feature that affects $\overline{\Delta RS}_m$ is the importance of the merging parties' common circulation area to these two newspapers. This influences how strong the cross-effect is. Suppose that two newspapers only compete with each other in a county that is far away from their home counties. This county might not play a large role in generating profit for these two newspapers because of readers' taste for local newspapers. When this is the case, a change in the quality of one newspaper does not affect the profit of the other newspaper much and thus the cross-effect is weak. Hence, the post-merger adjustment is small. This feature is captured by the pre-merger overlapping rate of the two newspapers under ownership consolidation:

$$\text{overlap}_m = \left(\sum\nolimits_{c \in CTY_{1,2}} q_{1mc}\right) / q_{1m},$$

where $CTY_{1,2}$ is the intersection of the markets of the two newspapers, and q_{1mc} and q_{1m} are county circulation (in county c) and total circulation of the larger newspaper, respectively.[22] The above regression indicates a negative correlation between $\overline{\Delta RS}_m$ and overlap_m, meaning that the larger the overlapping area is, the larger is the welfare loss for readers.

[22] The pre-merger overlapping rate can be also defined for the smaller newspaper in consolidation as $\left(\sum_{c \in CTY_{1,2}} q_{2mc}\right) / q_{2m}$, where q_{2mc} and q_{2m} are similarly defined for the second largest newspaper. It is not included in the regression, because it is 1 in the majority of the markets simulated.

The third factor is the pre-merger asymmetry of the two parties to the merger in terms of circulation, measured by $\log\left(\frac{q_{1m}}{q_{2m}}\right)$. As explained in Section 5, since the marginal benefit of increasing circulation is larger for a larger newspaper, the publisher of the merged parties typically increases the quality of its larger newspaper or downgrades the larger newspaper by a smaller margin than that of its smaller newspaper. As an adjustment in a larger newspaper has a bigger impact on readers' welfare than the same adjustment in a smaller newspaper, asymmetry matters. Specifically, the larger the asymmetry, the smaller the welfare loss for readers, as indicated by the positive sign in the above regression.

Finally, the presence of a competitor mitigates the welfare loss for readers. Therefore, the triopoly dummy has a positive sign in the regression. On average, the welfare loss of a merger in a triopoly market is 4.85 dollars less than that in a duopoly market.

The second regression studies the bias in welfare effect when quality adjustment is ignored. The regression result is as follows:

$$\Delta\overline{\Delta RS}_m = \underset{(2.87)}{3.76} - \underset{(1.19)}{2.10}\,\text{triopoly}_m + \underset{(6.19)}{12.61}\,\text{pen}_m - \underset{(5.23)}{14.18}\,s_{1m} + \varrho_m.$$

Again, the triopoly dummy and pre-merger penetration matter. For example, the positive coefficient of pen_m means that a higher penetration is related to a larger bias in measuring welfare change. Another factor that determines the bias is the demand elasticity with respect to price. This is because the bias is generated by the difference between the post-merger/without-quality-adjustment equilibrium and the post-merger/with-quality-adjustment equilibrium. The bias is therefore the welfare effect of a policy that forces a publisher to set newspaper quality back to the pre-merger level. Suppose the pre-merger quality of one newspaper is higher. Then, as the publisher increases the quality of the newspaper under this policy, it can increase the price. How much it will increase the price depends on the price elasticity of demand.

Since the price elasticity in a logit model positively depends on market shares when market shares are smaller than 1/2, I use s_{1m}, the pre-merger market penetration of the large newspaper in its largest circulation county, to capture this factor. The negative sign in the regression result implies that a higher price elasticity of demand is related to a smaller bias from ignoring quality adjustment.

7 Conclusion

In this paper, I set up a structural model of the U.S. daily newspaper market and study the welfare implications of newspaper ownership consolidation, taking into account endogenous product choice as well as price choices. A large new data set is collected to estimate the model. Based on the estimates, I study mergers in the Minneapolis market. I also quantify the welfare implications of ownership consolidation in all duopoly and triopoly markets in the 2005 sample. The distribution of the welfare effects across markets is used to study the correlation between the welfare effect of ownership consolidation in a market and the structure of the market. The main findings are as follows.

First, in the counterfactual ownership consolidation of the *Star Tribune* and the *St. Paul Pioneer Press* in the Minneapolis market, the publisher of these two newspapers enlarges product differentiation by increasing the overall quality of *Star* and decreasing that of *Pioneer*. Subscription prices of both newspapers increase. The overall change in characteristics and prices leads to a decrease in circulation of both newspapers, with a larger drop in *Pioneer*. The resulting welfare impacts on readers, advertisers and publishers are -4.02 million dollars, -4.49% of the pre-merger advertiser surplus and 15.03 million dollars, respectively.

Second, ignoring quality adjustment can lead to a bias in estimating the change in prices and circulation, and hence a bias in welfare effects. For example, in the ownership consolidation of the *Star Tribune* and the *St. Paul Pioneer Press*, the price adjustment is underestimated by 25% of the average newspaper price and the publisher surplus is underestimated by more than 1 million dollars when quality adjustment is ignored. Quality adjustment is potentially also important for other changes in market structure such as joint operation agreement or changes in trade policies. These are interesting topics for future research.

Third, the simulation results of mergers in duopoly and triopoly markets show that the existence of a competitor unsurprisingly mitigates the loss in readers' welfare due to a competition effect. Readers' welfare loss resulting from ownership consolidation in a market is positively correlated with how much households in the market care about newspapers in general and how important the overlapping area of the two merged parties is to these two newspapers. It is negatively correlated with the asymmetry of newspaper size measured by pre-merger circulations. The magnitude of these correlations is reported in the last section. This part of the study can be used to guide policy works.

References

Audit Bureau of Circulations, *Newspaper County Penetration Reports*, 1998-2006 edition. Schaumburg, IL.

Bacon's Information Inc., *Bacon's Newspaper Directory*, 1998-2006 edition. Chicago, IL.

Berry, Steven (1994), "Estimating discrete-choice models of product differentiation." *RAND Journal of Economics*, 25, 242–262.

Berry, Steven, James Levinsohn, and Ariel Pakes (1995), "Automobile prices in market equilibrium." *Econometrica*, 63, 841–90.

Berry, Steven and Joel Waldfogel (2001), "Do mergers increase product variety? Evidence from radio broadcasting." *The Quarterly Journal of Economics*, 116, 1009–1025.

Berry, Steven and Joel Waldfogel (2003), "Product quality and market size." NBER Working Paper 9675, National Bureau of Economic Research.

Chandra, Ambarish and Allan Collard-Wexler (2009), "Mergers in two-sided markets: An application to the Canadian newspaper industry." *Journal of Economics & Management Strategy*, 18, 1045–1070.

Chu, Chenghuan Sean (2008), "The effect of satellite entry on product quality for cable television." Finance and Economics Discussion Series 2008-12, Board of Governors of the Federal Reserve System.

Crawford, Gregory S. and Matthew Shum (2006), "The welfare effects of endogenous quality choice: The case of cable television." mimeo, University of Arizona.

Crawford, Gregory S. and Ali Yurukoglu (2009), "The welfare effects of bundling in multi-channel television markets." Working Paper, University of Warwick.

Dertouzos, James N. and William B. Trautman (1990), "Economic effects of media concentration: Estimates from a model of the newspaper firm." *The Journal of Industrial Economics*, 39, 1–14.

Draganska, Michaela, Michael Mazzeo, and Katja Seim (2009), "Beyond plain vanilla: Modeling joint product assortment and pricing decisions." *Quantitative Marketing and Economics*, 7, 105–146.

Editor & Publisher, *Editor & Publisher International Year Book*, 1998-2006 edition. New York, NY.

Eizenberg, Alon (2009), "Upstream innovation and product variety in the U.S. home PC market." Working Paper, University of Pennsylvania.

Ferguson, James M (1983), "Daily newspaper advertising rates, local media cross-ownership, newspaper chains, and media competition." *Journal of Law & Economics*, 26, 635–54.

Gandhi, Amit, Luke Froeb, Steven Tschantz, and Gregory Werden (2008), "Post-Merger Product Repositioning." *The Journal of Industrial Economics*, 56, 49–67.

Genesove, David (1999), "The adoption of offset presses in the daily newspaper industry in the United States." NBER Working Paper 7076, National Bureau of Economic Research.

Gentzkow, Matthew (2007), "Valuing new goods in a model with complementarity: Online newspapers." *American Economic Review*, 97, 713–744.

Gentzkow, Matthew and Jesse Shapiro (2010), "What drives media slant? Evidence from U.S. daily newspapers." *Econometrica*, 35–71.

George, Lisa (2007), "What's fit to print: The effect of ownership concentration on product variety in daily newspaper markets." *Information Economics and Policy*, 19, 285–303.

Hendel, Igal (1999), "Estimating multiple-discrete choice models: An application to computerization returns." *The Review of Economic Studies*, 66, 423–446.

Knight, Brian G. and Chun-Fang Chiang (2008), "Media bias and influence: Evidence from newspaper endorsements." NBER Working Paper 14445, National Bureau of Economic Research.

Lustig, Josh (2009), "Measuring welfare losses from adverse selection and imperfect competition in privatized medicare." Working Paper, Boston University.

Mazzeo, Michael (2002), "Product choice and oligopoly market structure." *The RAND Journal of Economics*, 33, 221–242.

Nevo, Aviv (2000), "Mergers with differentiated products: The case of the ready-to-eat cereal industry." *RAND Journal of Economics*, 31, 395–421.

Nevo, Aviv, Daniel L. Rubinfeld, and Mark McCabe (2005), "Academic journal pricing and the demand of libraries." *American Economic Review*, 95, 447–452.

Rosse, James (1967), "Daily newspapers, monopolistic competition, and economies of scale." *The American Economic Review*, 57, 522–533.

Rosse, James N. (1970), "Estimating cost function parameters without using cost data: Illustrated methodology." *Econometrica*, 38, 256–275.

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

Schulhofer-Wohl, Sam and Miguel Garrido (2009), "Do newspapers matter? Evidence from the closure of the Cincinnati Post." Working Paper 1140, Princeton University, Woodrow Wilson School of Public and International Affairs, Discussion Papers in Economics.

Seim, Katja (2006), "An empirical model of firm entry with endogenous product-type choices." *RAND Journal of Economics*, 37, 619–640.

Small, Kenneth and Harvey Rosen (1981), "Applied welfare economics with discrete choice models." *Econometrica*, 49, 105–130.

Sonnac, Nathalie (2000), "Readers' attitudes toward press advertising: Are they ad-lovers or ad-averse?" 13, 249–259.

SRDS, *SRDS Circulation*, 1998-2006 edition. Des Plaines, IL.

Sweeting, Andrew (2007), "Dynamic product repositioning in differentiated product markets: The case of format switching in the commercial radio industry." NBER Working Paper 13522, National Bureau of Economic Research.

Watson, Randal (2009), "Product variety and competition in the retail market for eyeglasses." *Journal of Industrial Economics*, 57, 217–251.

Appendix

A Data Sources and Definition of Variables

Demand. Data on county circulation for newspapers that are members of the Audit Bureau of Circulations (ABC) is from the *Newspaper County Circulation Report.* ABC members account for about 2/3 of all daily newspapers in the U.S. For non-ABC members, county circulation figures are from newspapers' sworn postal statements available in *SRDS Circulation.* Display advertising linage data is available for 485 newspaper/years between 1999 and 2005. The data comes from *TNS Media Intelligence.*

Prices. Data on annual subscription prices and display advertising rates is from *Editor and Publisher International Year Book (E&P).* A tiny number of newspapers have multiple subscription prices. The local price is used. Display advertising rate is the open inch rate measured in dollars per column inch.

Characteristics. Data on average pages per issue is from *E&P.* It is defined as the weighted sum of average pages per issue for weekdays and that for Sunday with weights $(\frac{6}{7}, \frac{1}{7})$.

The number of opinion section staff and reporters is collected from *Bacon's Newspaper Directory.* Bacon's Directory provides information on the titles, for example "Business Reporter", and names of all managing and editorial staff for all daily newspapers in the U.S. For each newspaper, I collect the name of all reporters and assign a weight to each one of them. The weight is the inverse of the number of titles that a reporter has. I then sum up the weights to get "the number of reporters". For example, if a person is a reporter and has only one title, she is counted as 1. If she is a court reporter and a crime reporter, she is counted as 1 as well. But if she holds some managing job at the same time and has therefore another entry in the directory, she contributes to 2/3 in "the number of reporters". The number of columnists and editorial editors is similarly defined.

Data on frequency of publication and edition (morning or evening newspaper) is from *E&P.* So is the information on the home county of a newspaper. The distance of two counties is computed based on the data of latitude and longitude of county centers provided by the Census Bureau.

County Demographics. The number of households is from the Audit Bureau of Circulations. Other county demographics are from the 2000 Census.

The data source and the description of the variables are summarized in Table 10.

Table 10: Data Description and Source

	Variable	Data Description	Data Source
Newspaper Demand	q_{jct}	County circulation	ABC, SRDS
Display Advertising Demand	a_{jt}	Annual display advertising linage (column inch)	TNS
Price of Newspaper	p_{jt}	Annual subscription price (1997 $)	E&P
Price of Display Advertising	r_{jt}	Adverting rate (1997 $/column inch)	E&P
Newspaper Characteristics	x_{2jt}	Weighted sum of reporters and correspondents	Bacon
	x_{3jt}	Weighted sum of columnists and editorial editors	Bacon
	f_{jt}	Frequency of publication (issues/52 weeks)	E&P
	y_{2jt}	Edition (morning or evening)	E&P
	n_{jt}	Average pages per issue	E&P
County Distance	y_{4jct}	Distance between county c and the home county of newspaper j	E&P, Census
Ownership		Publisher	Bacon
County Demographics	z_{1c}	% of population over 25 with bachelor's degree or higher	Census
	z_{2c}	Median income (1997 $)	Census
	z_{3c}	Median age	Census
	z_{4c}	% of urban population	Census
	H_{ct}	Number of households	ABC

ABC: County Circulation Report by Audit Bureau of Circulations

Bacon: Bacon's Newspaper Directory

E&P: Editor and Publisher International Year Book

SRDS: SRDS Circulation

TNS: TNS Media Intelligence

B Welfare Measures

Reader Surplus

The compensating variation for household i is given by

$$CV_{ict} = \frac{V^0_{ict} - V^1_{ict}}{\alpha_i},$$

where $\alpha_i < 0$ is the negative of the household's marginal value of income, and $V^0_{ict} - \alpha_i I_i$ and $V^1_{ict} - \alpha_i I_i$ are the expected maximum utility for household i with income I_i (expectation with respect to the extreme value taste shocks) before and after a merger. Specifically,[23]

$$V^0_{ict} = \ln\left(\sum_{j=1}^{J_{ct}} e^{U^0_{ijct}} + 1\right) + \sum_{j=1}^{J_{ct}} \ln\left(\sum_{h\neq j} e^{U^0_{ihct}-\kappa} + 1\right) - (J_{ct}-1)\ln\left(\sum_{h=1}^{J_{ct}} e^{U^0_{ihct}-\kappa} + 1\right),$$

where $U^0_{ijct} = u^0_{ijct} - \varepsilon_{ijct}$ is the utility before the merger, net of the extreme value taste shock. The after-merger utility V^1_{ict} is analogously defined to V^0_{ict}, replacing U^0_{ijct} by U^1_{ijct} and u^0_{ijct} by u^1_{ijct}.

Given the compensating variation for a specific household above, change in the average per-household reader surplus in county c in year t is given by $\overline{\Delta RS_{ct}} = E_{\boldsymbol{\zeta}_i}(CV_{ict})$. Total welfare change is then the sum of the welfare change in all the counties in a game: $\Delta RS = \sum_{ct} H_{ct}\overline{\Delta RS_{ct}}$, where H_{ct} is the number of households in county c in year t; and change in average per-household reader surplus is $\overline{\Delta RS} = \frac{\Delta RS}{\sum_{ct} H_{ct}}$.

Advertiser Surplus

As mentioned in Section 5, I only have information to measure the percentage change in advertising surplus. This can be seen as follows. As in Rysman (2004), suppose a representative advertiser has the following maximization problem:

$$\max_{\{a_j\}} \sum_j \left(\eta_j' q_j^{\lambda_1'} A_j^{\lambda_2'} a_j^{\lambda_3'} - r_j a_j\right), 0 < \lambda_3' < 1, \eta_j' > 0, \tag{B.1}$$

where a_j is the advertising space that the advertiser purchases in newspaper j and A_j is the total advertising space in newspaper j. The total advertising space influences the visibility of a specific advertisement. When λ_2' is negative, there exist negative externalities in advertising.

The solution to the advertiser's problem is

$$a_j = \left(\lambda_3'\eta_j'\right)^{\frac{1}{1-\lambda_3'}} q_j^{\frac{\lambda_1'}{1-\lambda_3'}} A_j^{\frac{\lambda_2'}{1-\lambda_3'}} r_j^{\frac{1}{\lambda_3'-1}}. \tag{B.2}$$

Aggregation (setting $a_j = A_j$) yields

$$A_j = \left(\lambda_3'\eta_j'\right)^{\frac{1}{1-\lambda_2'-\lambda_3'}} q_j^{\frac{\lambda_1'}{1-\lambda_2'-\lambda_3'}} r_j^{\frac{1}{\lambda_2'+\lambda_3'-1}},$$

[23] The derivation of this expression follows directly from Small and Rosen (1981) for a single discrete choice model. The only difference is the second and third term, the sum of which is the expectation (with respect to the extreme value taste shocks) of the second highest utility.

which can be rewritten as follows with $\lambda_1 = \frac{\lambda_1'}{1-\lambda_2'-\lambda_3'}$, $\lambda_2 = \frac{1}{\lambda_2'+\lambda_3'-1}$ and $\eta_j = \log\left[\left(\lambda_3'\eta_j'\right)^{\frac{1}{1-\lambda_2'-\lambda_3'}}\right]$:

$$a\left(r_j, q_j, \eta_j; \boldsymbol{\lambda}\right) = e^{\eta_j} q_j^{\lambda_1} r_j^{\lambda_2}. \tag{B.3}$$

This is the advertising demand function in (6). Plugging it into the advertiser's profit function in (B.1) gives the measure for advertiser surplus

$$AS = \left(\frac{1}{\lambda_3'} - 1\right) a_j r_j,$$

where $\frac{1}{\lambda_3'-1}$ is the representative advertiser's demand elasticity with respect to price (see (B.2)). Since the representative advertiser's price elasticity parameter, λ_3', and the externality parameter, λ_2', cannot be identified separately given only aggregate data, I report the percentage change in advertiser surplus.

C Invertibility of the Penetration Function

In this appendix, I show that the invertibility result in BLP can be extended to a multiple discrete choice model. I only show the extension for a model where the number of products that an individual can buy is limited to at most two. The result can be easily extended to a model in which consumers can choose up to $\bar{n} \leq J$ products, where J is the total number of products available in a choice set.

Penetration Function

The penetration function in Section 2.1 is given by

$$\begin{aligned}
\mathscr{I}_j\left(\boldsymbol{\delta}, \boldsymbol{p}, \boldsymbol{x}; \boldsymbol{\sigma}, \kappa\right) &= \int \Psi_j^{(1)}\left(\boldsymbol{\delta}, \boldsymbol{p}, \boldsymbol{x}, \varsigma_i; \boldsymbol{\sigma}\right) d\Phi\left(\varsigma_i\right) \\
&\quad + \sum_{j' \neq j} \int \left(\Psi_{j,j'}^{(2)}\left(\boldsymbol{\delta}, \boldsymbol{p}, \boldsymbol{x}, \varsigma_i; \boldsymbol{\sigma}, \kappa\right) - \Psi_j^{(3)}\left(\boldsymbol{\delta}, \boldsymbol{p}, \boldsymbol{x}, \varsigma_i; \boldsymbol{\sigma}, \kappa\right)\right) d\Phi\left(\varsigma_i\right),
\end{aligned}$$

where

$$\Psi_j^{(1)}\left(\boldsymbol{\delta}, \boldsymbol{p}, \boldsymbol{x}, \varsigma_i; \boldsymbol{\sigma}\right) = \frac{\exp\left(\delta_j + \vartheta_{ij}\right)}{1 + \sum_{h=1}^{J} \exp\left(\delta_h + \vartheta_{ih}\right)},$$

is the probability that newspaper j is chosen as the first newspaper ($\vartheta_{ij} = \sigma_0 p_j \varsigma_{0i} + \sum_{k=1}^{3} \sigma_k x_{kj} \varsigma_{ki}$ is the deviation from the mean utility), and the probability that newspaper j is chosen as the second newspaper when j' is the first best is given by the difference between the followings:

$$\begin{aligned}
\Psi_{j,j'}^{(2)}\left(\boldsymbol{\delta}, \boldsymbol{p}, \boldsymbol{x}, \varsigma_i; \boldsymbol{\sigma}, \kappa\right) &= \frac{\exp\left(\delta_j + \vartheta_{ij}\right)}{\exp\left(\kappa\right) + \sum_{h \neq j'} \exp\left(\delta_h + \vartheta_{ih}\right)}, \\
\Psi_j^{(3)}\left(\boldsymbol{\delta}, \boldsymbol{p}, \boldsymbol{x}, \varsigma_i; \boldsymbol{\sigma}, \kappa\right) &= \frac{\exp\left(\delta_j + \vartheta_{ij}\right)}{\exp\left(\kappa\right) + \sum_{h=1}^{J} \exp\left(\delta_h + \vartheta_{ih}\right)}.
\end{aligned}$$

Invertibility

Since all statements in this section are true for any given $(\boldsymbol{p}, \boldsymbol{x}, \boldsymbol{\sigma}, \kappa)$, these arguments in $\mathscr{s}_j$ are omitted for presentational simplicity.

The proof of the invertibility result is slightly different from that in BLP. BLP define a function $F : R^J \to R^J$ pointwise as $F_j(\boldsymbol{\delta}) = \delta_j + \ln s_j - \ln \mathscr{s}_j(\boldsymbol{\delta})$ and show that F is a contraction when an upper bound on the value taken by F is imposed. For a single discrete choice model, the value of δ_j that solves $\sum_{h=1}^J s_h = \sum_{h=1}^J \mathscr{s}_h(\boldsymbol{\delta})$ when $\delta_{j'} = -\infty$ for $\forall j' \neq j$ is the upper bound of the jth dimension of a fixed point of F. In a multiple discrete choice model, however, this value does not exist when $\left(\sum_{h=1}^J s_h\right)$ is larger than 1.[24]

I first prove the existence and uniqueness of the solution to $\mathscr{s}_j(\boldsymbol{\delta}, \boldsymbol{p}, \boldsymbol{x}; \boldsymbol{\sigma}, \kappa) = s_j$ for all j directly without using the function F. I then verify that all conditions in BLP hold so that F is indeed a contraction mapping – when an upper bound is imposed.

The following inequalities, which will be shown at the end of this section, are useful in the proof:

$$\begin{aligned} \partial\mathscr{s}_j/\partial\delta_j &< \mathscr{s}_j && \text{(C.1)} \\ \partial\mathscr{s}_j/\partial\delta_j &> 0 && \text{(C.2)} \\ \partial\mathscr{s}_j/\partial\delta_h &< 0 \text{ when } h \neq j && \text{(C.3)} \\ \sum\nolimits_{h=1}^J (\partial\mathscr{s}_j/\partial\delta_h) &> 0 && \text{(C.4)} \end{aligned}$$

Inequalities (C.2), (C.3) and (C.4) imply that the Jacobian of $\mathscr{s}$ has a dominant diagonal. Therefore, there is a unique solution to the equation system of $\mathscr{s}_j(\boldsymbol{\delta}) = s_j$ for all j.[25]

I now show that all conditions in the theorem in BLP hold.

Condition (1): Inequalities (C.1) and (C.3) imply that

$$\begin{aligned} \partial F_j(\boldsymbol{\delta})/\partial\delta_j = &\; 1 - (\partial\mathscr{s}_j/\partial\delta_j)/\mathscr{s}_j > 0 \\ \partial F_j(\boldsymbol{\delta})/\partial\delta_h = &\; -(\partial\mathscr{s}_j/\partial\delta_h)/\mathscr{s}_j > 0 \text{ when } h \neq j. \end{aligned}$$

Also, inequality (C.4) implies that

$$\sum\nolimits_{h=1}^J \partial F_j(\boldsymbol{\delta})/\partial\delta_h = 1 - \sum\nolimits_{h=1}^J (\partial\mathscr{s}_j/\partial\delta_h)/\mathscr{s}_j < 1.$$

Condition (2): Given the monotonicity of F in all dimensions of $\boldsymbol{\delta}$, a lower bound of function F is $\underline{\delta} = \min_j \left(\lim_{\boldsymbol{\delta}\to-\infty^J} F_j(\boldsymbol{\delta})\right)$.

[24] In a single discrete choice model, $\sum_{h=1}^J s_h < 1$, while in a multiple discrete choice model, the sum of market penetration for all products $\sum_{h=1}^J s_h$ can be larger than 1. But the supremum of $\sum_{h=1}^J \mathscr{s}_h(\boldsymbol{\delta})$ is 1 when $\delta_{j'} = -\infty$ for $\forall j' \neq j$.

[25] See McKenzie, Lionel (1959), "Matrices with dominant diagonals and economic theory." In *Mathematical methods in the social sciences* (Kenneth Joseph Arrow, Samuel Karlin, and Patrick Suppes, eds.), 47-62, Stanford University Press.

Condition (3): I have already shown that the equation system of $\mathscr{s}_j(\boldsymbol{\delta}) = s_j$ has a unique solution. This implies that the mapping F has a unique fixed point. Denote the fixed point by $\boldsymbol{\delta}^*$. Then, $F_j(\boldsymbol{\delta}^*) = \delta_j^*$ for all j. Note that $F_j(\boldsymbol{\delta}^* + \Delta) - (\delta_j^* + \Delta) = \ln s_j - \ln \mathscr{s}_j(\boldsymbol{\delta}^* + \Delta)$ is strictly decreasing in Δ as implied by inequality (C.4). Therefore, $F_j(\boldsymbol{\delta}^* + \Delta) < (\delta_j^* + \Delta)$ for any $\Delta > 0$. Define $\bar{\delta}_j = \delta_j^* + \Delta$. Then, $F_j(\bar{\boldsymbol{\delta}}) < \bar{\delta}_j$ for any j. By inequality (C.3), $F_j(\boldsymbol{\delta}) < \delta_j$ for any $\boldsymbol{\delta}$ such that $\delta_j = \bar{\delta}_j$ and $\delta_{j'} \leq \bar{\delta}_{j'}$ for all j'.

I now show inequalities (C.1) to (C.4). Three observations are important:

$$0 < \Psi_j^{(1)}, \Psi_{j,j'}^{(2)}, \Psi_j^{(3)} < 1;\ \Psi_{j,j'}^{(2)} > \Psi_j^{(3)};\ \Psi_j^{(1)} > \Psi_j^{(3)}.$$

Inequalities (C.1) and (C.3) follow directly from the three observations:

$$\begin{aligned}
\partial \mathscr{s}_j / \partial \delta_j &= \int \Psi_j^{(1)} \left(1 - \Psi_j^{(1)}\right) d\Phi(\boldsymbol{\varsigma}) + \sum_{j' \neq j} \int \left[\Psi_{j,j'}^{(2)} \left(1 - \Psi_{j,j'}^{(2)}\right) - \Psi_j^{(3)} \left(1 - \Psi_j^{(3)}\right)\right] d\Phi(\boldsymbol{\varsigma}) \\
&< \int \Psi_j^{(1)} d\Phi(\boldsymbol{\varsigma}) + \sum_{j' \neq j} \int \left(\Psi_{j,j'}^{(2)} - \Psi_j^{(3)}\right) d\Phi(\boldsymbol{\varsigma}) = \mathscr{s}_j, \\
\partial \mathscr{s}_j / \partial \delta_h &= -\int \Psi_j^{(1)} \Psi_h^{(1)} d\Phi(\boldsymbol{\varsigma}) + \int \Psi_j^{(3)} \Psi_h^{(3)} d\Phi(\boldsymbol{\varsigma}) + \sum_{j' \neq j,h} \int \left(-\Psi_{j,j'}^{(2)} \Psi_{h,j'}^{(2)} + \Psi_j^{(3)} \Psi_h^{(3)}\right) d\Phi(\boldsymbol{\varsigma}) \\
&< \sum_{j' \neq j,h} \int \left(-\Psi_{j,j'}^{(2)} \Psi_{h,j'}^{(2)} + \Psi_j^{(3)} \Psi_h^{(3)}\right) d\Phi(\boldsymbol{\varsigma}) < 0 \text{ when } h \neq j.
\end{aligned}$$

To show inequality (C.4), note that $\sum_{h=1}^{J} \frac{\partial \mathscr{s}_j(\boldsymbol{\delta})}{\partial \delta_h} = \frac{\partial \mathscr{s}_j(\boldsymbol{\delta}+\Delta)}{\partial \Delta}|_{\Delta=0}$, and

$$\begin{aligned}
&\frac{\partial \mathscr{s}_j(\boldsymbol{\delta} + \Delta)}{\partial \Delta}|_{\Delta=0} \\
= &\int \left(\Psi_j^{(1)}\right)^2 \frac{1}{e^{\delta_j + p_j \sigma_0 \varsigma_{0i} + \boldsymbol{x}_j \boldsymbol{\sigma} \boldsymbol{\varsigma}}} d\Phi(\boldsymbol{\varsigma}) + \sum_{j' \neq j,0} \int \left[\left(\Psi_{j,j'}^{(2)}\right)^2 - \left(\Psi_j^{(3)}\right)^2\right] \frac{e^{\kappa}}{e^{\delta_j + p_j \sigma_0 \varsigma_{0i} + \boldsymbol{x}_j \boldsymbol{\sigma} \boldsymbol{\varsigma}}} d\Phi(\boldsymbol{\varsigma}) > 0.
\end{aligned}$$

Combining inequalities (C.3) and (C.4) yields inequality (C.2).