Market Structure and Product Quality in the U.S. Daily Newspaper Market
(Job Market Paper)

Ying Fan†

November 29, 2008

Abstract

This paper studies the effects of market structure changes on newspaper quality, subscription prices and advertising rates. It provides a framework to quantify the welfare effect of ownership consolidations taking into account endogenous quality choice. I develop a structural model that captures key features of the U.S. daily newspaper market and propose an estimation strategy that allows me to study product choice with continuous characteristics in an oligopoly market. A new data set on the U.S. newspaper market is collected to identify the demand for newspapers, demand for advertising and the cost structure of newspaper production. Two sets of counterfactual simulation exercises are conducted. The first is a case study of a merger of two newspapers in the Minneapolis market that was blocked by the Department of Justice. The simulation suggests that if it were allowed, readers’ welfare would have declined by 6 dollars per household on average, and 15 dollars in the county that would have been affected most adversely. The second exercise quantifies the welfare implications of ownership consolidations in duopoly and triopoly markets. The median loss in reader surplus in duopoly mergers is 16 dollars per household and 5 dollars in triopoly mergers.

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

JEL Classification: L0, L1, L2, L8, M3
1 Introduction

This paper studies how ownership consolidation affects product quality and welfare in the U.S. daily newspaper market. The last 25 years witnessed an overall decline of 55% in the number of independently owned daily newspapers. The decline occurred in almost every state and cut across all circulation sizes. What are the consequences of such a change in market structure on newspaper quality? Specifically, will newspaper publishers increase or decrease the space devoted to news? Will they enlarge or shrink the opinion-oriented section of a newspaper? Will they provide more staff-written stories or utilize more material from news agencies? How will, concomitantly, the newspaper price and the advertising rate change? What are the implications for welfare? A multi-product firm on the one hand does not want its products to compete with each other (direct business stealing effect), but on the other hand does not want to leave space for competitors to move their products there. A similar intuition applies to the daily newspaper market. Even though few counties have two daily newspapers of the same publisher competing with each other, around 70% of newspaper companies own newspapers that compete with common competitors, through which a business stealing effect may be transmitted indirectly.

This paper provides a framework to study empirically how market structure affects newspaper quality and prices. I set up a structural model to capture three key features of the U.S. daily newspaper market. First, a newspaper publisher’s revenue comes from both selling newspapers and selling advertising space. The demand for advertising depends on the number of readers. Therefore, product choice and the newspaper price not only directly affect circulation revenue, but also indirectly affect advertising revenue. Second, a household may subscribe to more than one newspaper, which requires a multiple discrete choice model on the demand side. Third, since not only prices (i.e. the newspaper subscription price and the advertising rate), but also characteristics of newspapers are chosen by firms, I use a two-stage game, where newspaper publishers choose characteristics in the first stage and prices in the second stage.

Methodologically, incorporating these features of the newspaper market into the empirical model requires several extensions to the existing estimation techniques. On the demand side, I use a model of multiple discrete choices based on Hendel (1999) with three differences: I allow for diminishing utility from the second choice; the model ensures that a household only buys one copy of a newspaper; and micro data is not needed for estimation. I thus show that the Berry, Levinsohn and Pakes (1995) (henceforth, BLP) methodology can be generalized to this context.

---

1 Other examples in this literature are Rosse (1967), Ferguson (1983), Genesove (1999), Gentzkow and Shapiro (2006) and George (2007), of which George (2007) is most closely related to this paper. She also studies market structure and product differentiation in the daily 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 take the stance of modeling it explicitly.

2 See Rysman’s (2004) study of the Yellow Pages market for a similar setup.

3 For instance, in 84 county/year samples, total newspaper circulation is larger than the number of households in the county.

4 Other examples in this literature are Nevo, Rubinfeld and McCabe (2005) and Gentzkow (2007).
On the supply side, I allow for both prices and product quality to be endogenous. Standard merger analysis studies price effects only, ignoring changes in product quality. As will be shown, this is an important omission because newspapers make substantial quality adjustment after merger. However, endogenizing product choice introduces new computational challenges. In particular, players in the first-stage decision take into account the impact of product choice on the equilibrium price in the second stage. But computing equilibrium prices for each possible product choice is burdensome. I overcome this, using the observation that it is sufficient to know the gradient of the equilibrium price function at the data points to formulate the optimality conditions for the observed product characteristics. This gradient is obtained from the total derivative of the first order condition for prices. This approach allows me to develop a tractable estimation routine, whereas nesting an equilibrium solving procedure in an estimation algorithm is computationally prohibitive. This estimation strategy can be used in studying product choice problems in general. Existing papers in the literature either directly specify a profit function which is not derived from demand (such as Mazzeo (2002)), or focus on monopoly industries (such as Crawford and Shum (2006)), or markets with a naturally finite and discrete product choice set (such as Draganska, Mazzeo and Seim (2007)).

I collect a large new data set of newspaper characteristics, subscription prices, advertising rates, circulation and advertising linage for all U.S. daily newspapers between 1997 and 2005. Based on the estimates of the model, I analyze a counterfactual merger of two newspapers in the Minneapolis market that was blocked by the Department of Justice. The simulation results show that under the merger readers’ welfare in the market would have declined by 6 dollars per household on average, and 15 dollars in the county that would have been affected the worst. This welfare loss is generated by a combination of increased subscription prices (by 10% on average) and reduced opinion section staff as well as reporters in the smaller party of the merger, by 5% and 6% respectively.

I also quantify the welfare implications of ownership consolidation in duopoly markets and triopoly markets where data permits. The simulation results show that the median loss in readers’ surplus in duopoly mergers is 16 dollars per household and in triopoly mergers 5 dollars. In 94% of the markets simulated, total welfare unambiguously falls. When quality adjustment is ignored, the loss in readers’ welfare is typically underestimated. The median bias is 4 dollars per household in duopoly mergers and 2 dollars in triopoly mergers. The distribution of the welfare effects across markets is also used to study the correlation between the welfare effect of ownership consolidation in a market and its underlying structure.

The rest of the paper is organized as follows. Section 2 presents results from difference-in-difference estimates. A structural model is presented in Section 3. Estimation equations are also derived in this section. The data is described in Section 4 and the estimation is explained in Section 5. The estimation and simulation results are in Sections 6 and 7, respectively. Section 8 concludes.

---

5 Data on advertising linage is available for only a subset of newspaper/years.

6 In a triopoly merger, the publisher of the largest newspaper is assumed to buy the second largest.
2 Difference-in-difference Estimation

I have compiled a new data set that contains information on newspaper characteristics (such as the number of opinion section staff and the number of reporters), subscription prices, advertising rates, county circulation, and advertising linage for all U.S. daily newspapers between 1997 and 2005.\textsuperscript{7} Circulation data is at the newspaper/county/year level. Data on other variables are at the newspaper/year level. Details can be found in Section 4 and Appendix C. In the current section, I first show that the U.S. daily newspaper market is characterized by overlapping circulation areas. Thus, theoretically, an ownership consolidation of newspapers whose circulation areas overlap should have an impact on quality and price of newspapers. I then present results from difference-in-difference estimates and show that newspaper publishers indeed adjust the quality and prices of their newspapers after a change in market structure.

Table 1 shows that the U.S. daily newspaper market is not dominated by monopolists due to the partial overlapping of circulation areas. To measure the extent of the overlap of a newspaper pair, I compute the percentage of circulation in the common area as a fraction of the total circulation for each member of the pair. For example, for 6109 newspaper/year pairs in the data, the overlapping percentage is above 25% for both members.

<table>
<thead>
<tr>
<th>criterion</th>
<th>25%</th>
<th>20%</th>
<th>15%</th>
<th>10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of newspaper/year pairs</td>
<td>6109</td>
<td>6273</td>
<td>6692</td>
<td>7400</td>
</tr>
</tbody>
</table>

An ownership consolidation of two newspapers with overlap in their circulation area, in theory, should affect the quality and prices of these newspapers as their publisher internalizes the cross-effect between them. I now present results from difference-in-difference estimates. The results suggest that newspaper publishers indeed adjust the number of staff for opinion sections, the number of reporters, the subscription price and the advertising rate of their newspapers after a change in market structure.

For each of the above four variables, three difference-in-difference regressions are carried out, corresponding to three different market structure changes (from a newspaper’s point of view): type A, type B, type C. A newspaper is considered to experience a type A change in market structure in year $t$ if it is the object of a transaction, i.e. it is sold by one publisher to another in year $t-1$. A newspaper experiences a type B market structure change if (1) it competes directly or indirectly with a type A newspaper\textsuperscript{8} in year $t$, and (2) its publisher is also the new owner of a type A newspaper in year $t$. Type C market structure change applies to any other newspaper that satisfies criterion (1). These definitions are illustrated with the example in Figure 1, where newspapers are labeled so that newspaper A, B, C experience type A, B, C changes in market structure, respectively.

\textsuperscript{7}Advertising linage data is available for only 485 newspaper/years between 1999 and 2005.

\textsuperscript{8} Here, a type A newspaper is a newspaper who experiences a type A change in market structure.
In the example, newspapers A and C compete in county 1 directly and newspapers B and C are both in county 2. Therefore, newspapers A and B compete indirectly. Suppose their owners were originally publishers 1, 2, 3, respectively. When publisher 1 sells newspaper A to publisher 2, newspaper A, B, C experience type A, B, C changes in market structure, respectively.

Table 2: Difference-in-difference Estimates of the Effect of Market Structure Changes

<table>
<thead>
<tr>
<th></th>
<th>type A</th>
<th>type B</th>
<th>type C</th>
</tr>
</thead>
<tbody>
<tr>
<td>price</td>
<td>9.07(^a)</td>
<td>7.73</td>
<td>3.89</td>
</tr>
<tr>
<td>advertising rate</td>
<td>14.53</td>
<td>8.85</td>
<td>3.89</td>
</tr>
<tr>
<td>opinion staff</td>
<td>1.14</td>
<td>0.72</td>
<td>0.72</td>
</tr>
<tr>
<td>reporters</td>
<td>2.72</td>
<td>2.82</td>
<td>2.82</td>
</tr>
</tbody>
</table>

The difference-in-difference estimation results in Table 2 suggest that a change in market structure affects the quality and price choices of not only the newspaper under transaction, but also all other newspapers that compete directly or indirectly with this newspaper. The overall welfare implication of the transaction therefore depends on all the changes of all newspapers involved.

Note that the categorization of market structure changes used is very coarse. For example, the effect of the transaction in the above example would be different if C were owned by publisher 2 before and after the transaction, or if A were also competing with other newspapers in other counties, or even if newspaper B were also competing with other newspapers in other counties, etc. From this example, it is clear that market structure changes cannot be easily categorized sufficiently with simple indicators. Additionally, one goal of this paper is to study the welfare implications of ownership consolidations. A difference-in-difference estimation cannot provide an answer to this welfare question. In the remainder of the paper, I therefore set up and estimate a structural model of endogenous product quality choice.

\(^a\)Formally, the estimation equation is \(y_{jt} = T_t \alpha + X_j \beta + \gamma x_{jt} + \varepsilon_{jt}\), where \(y_{jt}\) is one of the four variables in the table, \(T_t = (1_{t=97}, ..., 1_{t=05})\), \(X_j = (x_{j98}, ..., x_{j05})\), \(x_{jt} = 1\) if newspaper \(j\) experiences a market structure change of a certain type in year \(t\), and \(x_{jt} = 0\) if it does not experience ANY market structure change in year \(t\). In other words, in the estimation for a type A market structure change, for example, only newspapers experiencing a type A change in market structure and newspapers not experiencing market structure change of any type are used. The estimate of \(\gamma\) is reported in the table.

\(^b\)Standard errors are in parentheses.
3 The Model

The profit of a newspaper comes from both selling newspapers to readers and selling advertising space to advertisers. In this section, I describe the demand for newspapers, the demand for advertising as well as the supply side of the model. Estimation equations are also derived.

A Brief Road Map

The demand on readers’ side is described with a multiple discrete choice model based on Hendel (1999) with three differences: I allow for diminishing utility from the second choice; the model ensures that a household only buys one copy of a newspaper; and the model can be estimated with aggregate data. Demand for newspapers is derived in equation (7) below.

Following Rysman (2004), I derive the demand for advertising from a representative advertiser’s decision. Advertising demand depends on circulation and the advertising rate of a newspaper as in equation (10).

The supply side is modeled as a complete information two-stage game, where newspaper publishers choose characteristics in the first stage and prices in the second stage. The two-stage structure is used to capture in a simple way that newspaper publishers have a longer decision horizon when they make quality decisions but change prices more often.\(^{11}\) The three basic elements of this game — the set of players, timing and information, and payoffs — are described in the three subsections of Section 3.3, respectively.

The revenue function is determined by the two demand systems. In Section 3.3.3, I then describe the cost structure: the marginal cost of increasing circulation (equation (12)), the marginal advertising sales cost (equation (13)) and the marginal cost of increasing a certain newspaper characteristic (equation (14)). Together with the two demand functions, they give the profit function that is relevant for the second-stage price decision (equation (15)), and the profit function that is relevant for the first-stage quality decision (equation in (16)).

From this, five estimation equations are derived. The first two are the model implications of the two demand systems ([S], [ADV]) and the latter three are the optimality conditions with respect to advertising rates [RFOC], subscription prices [PFOC], and newspaper characteristics [XFOC].

The detailed structure of the model follows. Throughout the paper, a symbol with tilde represents a function and the same symbol without tilde is the value of the function at some point.

3.1 Demand for Newspapers

The demand for newspapers is derived from the aggregation of heterogeneous households’ multiple discrete choices. 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 (\(\bar{n}\)) to 2.\(^{12}\)

---

\(^{11}\)Estimating a fully dynamic model is beyond the scope of this paper.

\(^{12}\)The model is generalizable to \(\bar{n} > 2\).
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 Jct. 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. The household first compares \( u_{ijct} \) for \( j = 0, 1, ..., Jct \) and chooses \( j^* \) with the highest utility. If it chooses to subscribe to some newspaper, it can then subscribe to a second. In particular, if \( j^* \neq 0 \), it then compares \( \max \{ u_{ijct} - \kappa, j = 1, ... j^* - 1, j^* + 1, ..., Jct \} \) to \( u_{i0ct} \) and again takes the choice with the highest utility. \( \kappa \) is a positive parameter that captures the diminishing utility from subscribing to a second newspaper. The probability of household \( i \) subscribing to newspaper \( j \) is therefore the sum of the probability of \( j \) being chosen as the first choice and that of \( j \) being chosen as the second choice:

\[
\Pr \left( u_{ijct} \geq \max_{h=0, ..., Jct} u_{ihct} \right) + \sum_{j \neq j, 0} \Pr \left( u_{ij'ct} \geq \max_{h=1, ..., Jct, h \neq j'} u_{ihct} \& u_{ijct} - \kappa \geq u_{i0ct} \right). \tag{1}
\]

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

\[
u_{ijct} = p_{jt} + x_{jt} \beta_i + y_{jt} \psi_i + D_{ct} \phi_i + \xi_{jct} + \varepsilon_{ijt}, \tag{2}
\]

where \( p_{jt} \) is the annual subscription price, and \( x_{jt} = (x_{1jt}, ..., x_{Kjt}) \) contains the endogenous newspaper characteristics that are chosen by the newspaper publishers in the model. They are news hole (the space of a newspaper devoted to news), the number of staff for opinion sections and the paper characteristics that are chosen by the newspaper publishers in the model. They are news hole (the space of a newspaper devoted to news), the number of staff for opinion sections and the paper characteristics that are assumed to be exogenous in the model because they rarely change over time. The variable \( y_{1jt} \) measures the overall size of the market of newspaper \( j \). Market size of a newspaper affects utility because, for example, 10 reporters covering a small region can write stories of different quality and in different quantities compared to 10 reporters serving a large area. \( y_{2jt} \) is an edition dummy with value 1 when newspaper \( j \) is a morning newspaper and 0 otherwise. \( y_{3jct} \) is the distance between

---

13 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 taste, which is operationalized by county-level demographics.

14 One way to understand why the market size of a newspaper affects utility is to think that both \( x_{kjt} \) and \( x_{kjt}/y_{2jt} \) (for example, both reporters and reporters per household) affect the quality of a newspaper. In the empirical implementation in Section 5, I specify the quality characteristics as \( \log(1 + x_{kjt}) \) and the overall size of the market of newspaper \( j \) as the logarithm of the number of households in its market, i.e. \( \log(y_{2jt}) \). The utility then depends on \( \sum_k \beta_{ki} \log(1 + x_{kjt}) + \psi_2 \log(y_{2jt}) \), which is equivalent to \( \sum_k \left[ (\beta_{ki} - \psi_{2b}) \log(1 + x_{kjt}) + \psi_{2b} \log \left( \frac{1 + x_{kjt}}{y_{2jt}} \right) \right] \), where \( \sum_k \psi_{2b} = \psi_2 \). The latter expression means that utility depends on, for example, reporters as well as reporters per household. The former expression, which is used in the model, is more flexible as it even allows utility to depend on \( (1 + x_{kjt})/y_{2jt} \) for \( b \neq 1 \). To see this, note that \( \sum_k \left[ (\beta_{ki} - b\psi_{2b}) \log(1 + x_{kjt}) + \psi_{2b} \log \left( \frac{1 + x_{kjt}}{y_{2jt}} \right) \right] = \sum_k \beta_{ki} \log(1 + x_{kjt}) + \psi_2 \log(y_{2jt}) \).

---
derive the market demand for newspapers, I define the "relative" county mean utility, \( \delta \), where the difference between the county mean utility from subscribing to newspaper \( j \) and the utility from the outside choice, where it captures some characteristics of the newspaper that are relevant for the readers but unobservable to econometricians and therefore not included in \( x_{jt} \) or \( y_{jct} \). It also captures a county-specific taste for newspapers that is not included in \( D_{ct} \). \( \varepsilon_{ijt} \) is an i.i.d. stochastic term representing an unobservable household specific taste. Finally, \( \beta_i = (\beta_{1i}, \ldots, \beta_{Ki}) \) where \( \beta_{ki} = \beta_k + \sigma_k \varsigma_{ki} \) is household \( i \)'s specific taste for the \( k \)th endogenous characteristic. \( \beta_k \) captures the mean taste, while \( \sigma_k \) captures heterogeneity in the taste for endogenous characteristics, households are assumed to have homogenous tastes for the exogenous characteristics \( y_{jct} \). \( \varsigma_{ki} \) has an identically and independently distributed standard normal distribution across characteristics and households. Let \( \Phi(\cdot) \) represent the distribution function of \( \varsigma_i = (\varsigma_{1i}, \ldots, \varsigma_{Ki}) \).

Instead of treating the utility from the outside good as fixed, I model it as a time trend, which is associated with the development of online news sources and the increase in internet penetration. Specifically, let

\[
u_{i0ct} = (t - t_0) \rho + \varepsilon_{i0t}
\]

be the utility from the outside choice, where \( t_0 \) is the first year in the data.

This concludes the description of the multiple discrete choice model. Aggregation follows. To derive the market demand for newspapers, I define the “relative” county mean utility, \( \delta_{jct} \), as the difference between the county mean utility from subscribing to newspaper \( j \) and that from the outside choice:

\[
\delta_{jct} = p_{jt} \alpha + x_{jt} \beta + y_{jct} \psi + D_{ct} \varphi + \xi_{jct} - (t - t_0) \rho.
\]

Let \( \vartheta_{ijt} = \sum_{k=1}^{K} \sigma_k x_{kjt} \varsigma_{ki} \). Then, \( u_{ijct} = [\delta_{jct} + (t - t_0) \rho] + \vartheta_{ijt} + \varepsilon_{ijt} \). The utility is now expressed as the sum of mean utility \( \delta_{jct} + (t - t_0) \rho \) and a deviation from the mean, \( \vartheta_{ijt} + \varepsilon_{ijt} \). Following the literature, I assume that \( \varepsilon_{ijt} \) is drawn from a type I extreme value distribution with location parameter 0 and scale parameter 1. Plugging (2), (3) and (4) into (1) yields the probability of a household with taste \( \varsigma_i \) choosing newspaper \( j \):

\[
\tilde{\Psi}_j^{(1)}(\delta_{ct}, x_{ct}, \varsigma_i; \sigma) + \sum_{j' \neq j}' \tilde{\Psi}_j^{(2)}(\delta_{ct}, x_{ct}, \varsigma_i; \sigma, \kappa) - \tilde{\Psi}_j^{(3)}(\delta_{ct}, x_{ct}, \varsigma_i; \sigma, \kappa),
\]

where \( \varsigma_i = (\varsigma_{1i}, \ldots, \varsigma_{Ki}) \), \( \sigma = (\sigma_1, \ldots, \sigma_K) \), \( \delta_{ct} = (\delta_{jct}, j = 1, \ldots, J_{ct}) \), \( x_{ct} = (x_{jt}, j = 1, \ldots, J_{ct}) \) and

\[
\tilde{\Psi}_j^{(1)}(\delta_{ct}, x_{ct}, \varsigma_i; \sigma) = \frac{\exp(\delta_{jct} + \vartheta_{ijt})}{1 + \sum_{h=1}^{J_{ct}} \exp(\delta_{hct} + \vartheta_{ iht})},
\]

\[
\tilde{\Psi}_j^{(2)}(\delta_{ct}, x_{ct}, \varsigma_i; \sigma, \kappa) = \frac{\exp(\delta_{jct} + \vartheta_{ijt})}{\exp(\kappa) + \sum_{h \neq j'} \exp(\delta_{hct} + \vartheta_{ iht})},
\]

\[
\tilde{\Psi}_j^{(3)}(\delta_{ct}, x_{ct}, \varsigma_i; \sigma, \kappa) = \frac{\exp(\delta_{jct} + \vartheta_{ijt})}{\exp(\kappa) + \sum_{h=1}^{J_{ct}} \exp(\delta_{hct} + \vartheta_{ iht})}.
\]
The first summand is the probability of newspaper $j$ being chosen as the first newspaper. $\tilde{\Psi}^{(2)}_{j,j'}$ and $\tilde{\Psi}^{(3)}_j$ are probabilities of it being chosen as the second newspaper when household $i$ faces a constrained choice set with newspaper $j'$ excluded and when household $i$ faces an unconstrained choice set, respectively. The difference of these two is the probability of newspaper $j$ being chosen as the second newspaper when $j'$ is the first best.

County market penetration\textsuperscript{15} is the aggregation of households’ newspaper choices in a county:

$$
\tilde{s}_{jct} (\delta_{ct}, x_{ct}; \sigma, \kappa) = \int \tilde{\Psi}^{(1)}_j d\Phi (\varsigma_i) + \sum_{j' \neq j, 0} \int \left( \tilde{\Psi}^{(2)}_{j,j'} - \tilde{\Psi}^{(3)}_j \right) d\Phi (\varsigma_i),
$$

(6)

If there is only one newspaper in county $c$, $\tilde{s}_{jc} (\delta_{ct}, x_{ct}; \sigma) = \int \tilde{\Psi}^{(1)}_j d\Phi (\varsigma_i)$ as in a single discrete choice model. Let the last element of the county demographics vector, $D_{Lct}$, be the number of households in county $c$ in year $t$. The demand for newspaper $j$, i.e. the total circulation of newspaper $j$, is then the sum of circulations in all counties covered by newspaper $j$ (denoted by $C_{jt}$)

$$
\tilde{q}_{jt} (\delta_{ct}, x_{ct}; \sigma, \kappa) = \sum_{c : c \in C_{jt}} D_{Lct} \tilde{s}_{jct} (\delta_{ct}, x_{ct}; \sigma, \kappa).
$$

(7)

Note that I have assumed that readers only care about news hole and do not care about advertising. This assumption is necessary to keep the model tractable. The demand for advertising depends on the total circulation, in other words, readers’ decisions. If readers’ decisions in turn depend on advertisers’ decisions, solving an equilibrium of subscription price and advertising rate becomes a fixed-point problem. Moreover, it is not clear how to separately identify the effect of news hole and that of advertising on newspaper demand. Any exogenous variation that leads to changes in news hole also affects the advertising demand through influencing circulations. Similarly, because advertising affects circulation and thus affects publishers’ optimal choice of news hole, any exogenous source of change in advertising will also alter news hole.\textsuperscript{16}

I now derive the first estimation equation that will be taken to the data. Following Berry (1994), I do not use the demand equation directly. Instead, I use relative mean utility in equation (4) to avoid a nonlinear endogeneity problem.\textsuperscript{17} Berry (1994) and BLP show that in a single discrete choice model, under certain regularity conditions on the density of households’ unobservable tastes, there exists a unique vector of mean utility levels, $\delta_{ct}$, such that $s_{jct} = \tilde{s}_{jct} (\delta_{ct}, x_{ct}; \sigma)$.\textsuperscript{18} Theorem 1 below states that this invertibility result can be extended to the current multiple discrete choice model. Furthermore, the contraction mapping defined in BLP is still viable, leading to a simple algorithm to solve for $\delta_{ct}$.

\textsuperscript{15}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.

\textsuperscript{16}Rysman (2004) allows consumers to value advertising in his study of the network effects in the Yellow Pages market. But for one thing, Yellow Pages directories are free. Publishers choose advertising rates only and do not have a two-dimensional interdependent price decision. For another, there is no analogue of news hole in Yellow Pages.

\textsuperscript{17}Berry (1994) notices that product prices (here product quality as well) are correlated with the taste shock $\xi_{jct}$, which is nonlinear the market penetration function, $\tilde{s}$. This therefore leads to a nonlinear endogeneity problem.

\textsuperscript{18}That is to say, $\delta_{ct}$ is uniquely determined by the data for given $\sigma$. One can therefore treat it as if it were observable. Note that the unobservable taste shock $\xi_{jct}$ is linear in $\delta_{jct}$. This becomes a conventional linear endogeneity problem.
Theorem 1 For any $(s, x) \in R^J \times R^{KJ}, \sigma \in R^K, \kappa \in R^+$ and distribution functions $P_\zeta(\cdot; \sigma)$, define operator $F : R^J \to R^J$ pointwise as $F_j(\delta) = \delta_j + \ln s_j - \ln \tilde{s}_j(\delta, x; P_\zeta, \sigma, \kappa)$, where

\[
\tilde{s}_j(\delta, x; P_\zeta, \sigma, \kappa) = \int \tilde{\Psi}^{(1)}_j dP_\zeta(\varsigma; \sigma) + \sum_{j' \neq j, 0} \int \int (\tilde{\Psi}^{(2)}_{j,j'} - \tilde{\Psi}^{(3)}_j) dP_\zeta(\varsigma; \sigma),
\]

and $\tilde{\Psi}^{(1)}_j(\delta, x, \varsigma; \sigma), \tilde{\Psi}^{(2)}_{j,j'}(\delta, x, \varsigma; \sigma, \kappa)$ and $\tilde{\Psi}^{(3)}_j(\delta, x, \varsigma; \sigma, \kappa)$ are defined in (5). If (1) $0 < s_j < 1$ for $\forall j = 1, \ldots, J$ and (2) $\sum_{j=1}^J s_j < 2$, then the operator $F$ has a unique fixed point.

The proof of the theorem can be found in Appendix A. The first assumption means that there is always some household choosing newspaper $j$ and some household not choosing it. The second assumption means that there is always some household with fewer than two newspapers. Under these two conditions, the solution to $s_{jct} = \tilde{s}_{jct}(\delta_{ct}, x_{ct}; \sigma, \kappa)$ is unique. Denote this solution by $\delta_{ct}(s_{ct}; \sigma, \kappa)$. Plugging it into the expression for the relative mean utility level in equation (4) gives for the true value of $(\alpha, \beta, \psi, \phi, \rho, \sigma, \kappa)$:

\[
\delta_{jct}(s_{ct}; \sigma, \kappa) = p_{jt} + x_{jt} + y_{jct} + D_{ct} - (t - t_0) \rho + \xi_{jct}, \forall jct.
\]

This is the first estimation equation. For the reader’s ease, I label all estimation equations with brackets. This equation is labeled as [S](8) because it is derived from the market penetration function $\tilde{s}_{jct}$. In the remainder of this section, subscript $t$ is omitted for ease of exposition and only restored in the estimation equations.

3.2 Demand for Advertising

Following Rysman (2004), I assume that a representative advertiser has the following maximization problem:

\[
\max_{\{a_j\}} \sum_j \left( \eta'_j q_j \lambda'_j a_j^\lambda_j A_j^{\lambda_j} - r_j a_j \right), 0 < \lambda'_j < 1, \eta'_j > 0,
\]

where $a_j$ is the advertising space in newspaper $j$, and $r_j$ and $q_j$ are the advertising rate and the total circulation of newspaper $j$, respectively. High circulation is expected to increase advertising effectiveness. $A_j$ is the total advertising space. It affects the visibility of a specific advertisement. When $\lambda'_j$ is negative, there exist negative externalities in advertising. $\eta'_j$ captures the demographics of newspaper $j$’s circulation area. It also influences the how effective advertising in newspaper $j$ is.

Additive separability of the profit function over newspapers means an advertisement in one newspaper is not a substitute or complement to advertisements in another newspaper in terms of generating profit. Then, the advertiser will keep advertising on each newspaper until the marginal profit from this newspaper is 0. In other words, it maximizes its profit by choosing advertisement $a_j$ for each newspaper $j$ separately.
The solution to the above problem is

\[ a_j = \left( \lambda_3 \eta_j \right)^{\frac{1}{1 - \lambda_3}} q_j^{\frac{\lambda_1}{1 - \lambda_3}} A_j^{\frac{\lambda_2}{1 - \lambda_3}} r_j^{\frac{1}{1 - \lambda_3}}. \]

Aggregation (setting \( a_j = A_j \)) yields

\[ A_j = \left( \lambda_3' \eta_j' \right)^{\frac{1}{1 - \lambda_3'}} q_j^{\frac{\lambda_1'}{1 - \lambda_3'}} A_j^{\frac{\lambda_2'}{1 - \lambda_3'}} r_j^{\frac{\lambda_3'}{1 - \lambda_3'}}. \]

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

\[ \tilde{a} (r_j, q_j, \eta_j; \lambda) = e^{\eta_j q_j^{\lambda_1} r_j^{\lambda_2}}. \]  

(10)

As mentioned, \( \eta_j \) captures the demographics of newspaper \( j \)'s circulation area. Specifically, I operationalize \( \eta_j \) as follows. Let \( D_c \phi \) be a linear combination of observable demographics of county \( c \). \( \eta_j \) is defined as the circulation-weighted sum of these county indices over the counties covered by newspaper \( j \): \( \eta_j = \sum_{c: c \in C_j} \frac{q_{jc}}{q_j} D_c \phi \). \( \phi \) is a vector of parameters to be estimated.

Let \( \iota_{jt} \) be an i.i.d. and mean zero measurement error for display advertising linage, then the second estimation equation is

\[ \log a_{jt} = \sum_{c: c \in C_{jt}} \frac{q_{jct}}{q_{jt}} D_c \phi + \lambda_1 \log q_{jt} + \lambda_2 \log r_{jt} + \iota_{jt}, \forall jt. \]  

[ADV](11)

### 3.3 Supply

Typically, the term “market” is 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: (1) all consumers in the area face the same choice set and (2) the suppliers of these choices in the area compete with each other and with no one else. However, in the daily newspaper industry, due to the partially overlapping circulation areas of newspapers, there is no geographic area that satisfies both criteria. For this reason, in the remainder of the paper, I use the term “choice set” for readers on the demand side and “the set of players” on the supply side. The latter term is justified because the supply side is modeled as a complete information two-stage game. I now specify the three basic elements of this game: the set of players, timing and information as well as payoffs.

#### 3.3.1 The set of players

When newspaper A and B compete in county 1 and newspaper B and C compete in county 2, newspaper A, B, C are all in one game because A and B, as well as B and C, are direct competitors, and hence A and C are competitor’s competitors. Therefore, due to the partial overlapping of newspaper coverage, all newspapers in the U.S. are potentially in one game. To limit the number
of players in a game, two assumptions are made. The first assumption is that a newspaper competes only with the newspapers in its Newspaper Designated Market. The Newspaper Designated Market (NDM) is a set of counties that a newspaper reports (to the Audit Bureau of Circulations, a nonprofit circulation-auditing organization, and advertisers) as the market it serves. It is a predetermined subset of the counties where a newspaper circulates. The second assumption is that the behavior of the three national newspapers Wall Street Journal, New York Times and USA Today is taken as given in the model.

An example in Figure 2 is used to illustrate the definition of a set of players. The formal definition follows. In the example, if a newspaper circulates in a county, it is in the oval representing this county. If this county is also in its NDM, the newspaper is shaded. For example, newspaper C circulates in county 2 and 3. But its NDM consists of county 2 only. Therefore, according to the first assumption, it only competes with B in county 2, and does not compete with newspaper D in county 3. Because A and B are direct competitors in county 1, B and C are direct competitors in county 2, and they do not have economic interaction with other newspapers, the set of players is given by the publishers of newspaper A, B and C.

Formally, two newspapers \( j \) and \( j' \) are defined as interacting directly if there exists at least one county that is in the NDMs of both newspapers. Two newspapers \( j \) and \( j' \) are defined as interacting if either \( j \) and \( j' \) interact directly or there exist a set of newspapers \( \{h_n\}_{1}^{N} \) such that \( j \) interacts with \( h_1 \) directly, \( h_n \) interacts with \( h_{n+1} \) directly for \( n = 1, \ldots, N - 1 \), and \( h_N \) interacts with \( j' \) directly. The set of players in a game is defined as the owners of the set of newspapers such that every newspaper interacts with some other newspaper in this set and none of the newspapers interacts with newspapers not in this set. In other words, a set of players is defined as the publishers of the closure of the interacting relation.\(^{19}\) In the rest of the paper, I refer to a newspaper in the closure as a player newspaper and its publisher as a player publisher.

Some more notation is necessary for the remaining description of a typical game. For a given game, let \( \mathcal{M} \) be the set of player publishers in the game with \( m \) being a typical element, \( \mathcal{J}_m \) be the set of player newspapers owned by \( m \) in this game. \( \mathcal{J} = \cup_{m \in \mathcal{M}} \mathcal{J}_m \) represents the set of player newspapers in the game.

\(^{19}\)This does not mean that all newspapers owned by a newspaper publisher are in one game. In fact, a newspaper publisher can be a player in different games. But since there is no economic interaction between two player newspapers in different games, one can without loss of generality label one newspaper publisher in two different games as two different newspaper publishers.
Because of the two assumptions above, there might be newspapers that circulate in the NDMs of the player newspapers in a game but are not player newspapers in the game. 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 choices of quality and prices do not depend on those of the players. In other words, their quality and prices are taken as given in the game.

### 3.3.2 Timing and Information

The timing of the game is illustrated in Figure 3. The set of newspapers that each newspaper publisher owns, the NDM for each newspaper and the county demographics are predetermined before the start of the game. The exogenous newspaper characteristics, \( y \), are predetermined as well. As explained above, all aspects of a non-player newspaper are taken as exogenous in the model. They are assumed to be realized before the start of the game.

At the beginning of the game, the shocks are revealed: the newspaper/county specific taste \((\xi_{jc})\) and the marginal cost shocks \((\nu_{kj}, \omega_j, \zeta_j)\) to be specified below. The information is public to all players in the game. Given this information, all players simultaneously choose quality characteristics of their newspapers in the first stage. In the second stage, all players observe the newspaper characteristics and choose prices, including newspaper subscription prices and advertising rates, simultaneously.

### 3.3.3 Payoffs

The profit of a newspaper publisher comes from both circulation profit and advertising profit. The advertising demand described in Section 3.2 is really demand for display advertising, which is printed on the newspapers’ pages along with the news. In fact, there exists another type of

---

20 But non-players quality and prices affect the players’ decisions as they influence the newspaper demand.

21 It is actually a newspaper/county/year specific shock. The subscript \(t\) is omitted here and also in \((\nu_{kj}, \omega_j, \zeta_j)\).
advertisement: preprint. Preprints are inserted into each copy of a newspaper and distributed along with it. This is essentially a delivery service provided by newspapers. I do not observe the advertising rate for preprint. Therefore, the preprint profit is not derived from a demand model. Instead, I assume that it is a simple quadratic function of circulation:

\[ \mu_1 q_j + \frac{1}{2} \mu_2 q_j^2. \]

I now specify the cost structure. The demand for newspapers described in Section 3.1 and the demand for display advertising in Section 3.2 are both for annual demand: annual subscribers and annual advertising linage. Correspondingly, the costs modeled below are annual costs.

The cost of a newspaper consists of two parts: variable cost (variable with production) and fixed cost (fixed with respect to production). One variable cost is the cost of printing and delivery. It varies with circulation, \( q_j \), and its marginal depends on publication frequency and the number of pages. I assume this marginal cost, \( mc_j(q) \), to be constant to circulation:

\[ mc_j(q) = \gamma_1 + \gamma_2 f_j + \gamma_3 n_j f_j + \omega_j, \]

where \( f_j \) is the publication frequency measured by the number of issues per year, \( n_j \) is the average pages per issue, and \( \omega_j \) is a shock to the marginal cost. The annual number of pages, \( n_j f_j \), is the sum of annual news hole \( (x_{1j}) \) and display advertising linage \( (a_j) \). Hence, the marginal cost can now be expressed in terms of characteristics, advertising linage and the cost shock:

\[ \tilde{mc}(f_j, x_{1j}, a_j, \omega_j; \gamma) = \gamma_1 + \gamma_2 f_j + \gamma_3 (x_{1j} + a_j) + \omega_j. \]

(12)

Note that no other newspaper characteristics besides news hole \( (x_{1j}) \) and frequency \( (f_j) \) affect the marginal cost. That is because the cost of increasing some characteristics of a newspaper, such as the number of reporters, is independent of circulation.

Another variable cost is the advertising sales cost. It is assumed to be

\[ \tilde{mc}^{(a)}(\zeta_j; \bar{\zeta}, \lambda_2) = (1 + 1/\lambda_2) (\bar{\zeta} + \zeta_j), \]

(13)

where \( \lambda_2 \) is the price elasticity of display advertising demand and \( \zeta_j \) is a mean-zero exogenous random variable.

Finally, the fixed cost (fixed with respect to circulation and advertising sales) consists of the cost of choosing certain combination of newspaper quality characteristics. I assume that the marginal cost of increasing the \( k^{th} \) endogenous characteristic \( x_{kj} \) is

\[ \tilde{mc}^{(x)}(x_{kj}, \nu_j; \tau_k) = \tau_{k0} + \tau_{k1} x_{kj} + \nu_{kj}, \]

(14)

---

22 In a more general specification, I do not find significant evidence of economies of scale.

23 There is a slight abuse of notation here. The annual display advertising linage, \( a_j \), is measured by column inches. According to Editor and Publisher International Year Book, a typical U.S. daily newspaper page has 6 columns with 21-inch depth. Therefore, in fact, \( n_j f_j = x_{1j} + \frac{a_j}{6} \).

24 \( \lambda_2 \) is added so that the optimal display advertising rate condition [RFOC](17) is simple.
where $\nu_{kj}$ is the shock to the marginal cost of increasing it. The fixed cost $\tilde{fc}(x_j, \nu_j; \tau)$ is then the sum of the integrals $(\sum^K_{k=1}(\tau_0 + \frac{1}{2}r_k x_{kj} + \nu_{kj})x_{kj})$ plus a constant.

Let $\theta = (\alpha, \beta, \psi, \varphi, \rho, \sigma, \kappa, \phi, \lambda, \mu, \gamma, \zeta)$ be the collection of parameters that are relevant for the second-stage decision. Denote the variable profit from newspaper $j$ by $\tilde{\pi}^H_j (p, r; x, y, D, \xi, \omega, \zeta; \theta)$, where $p$ is a vector of subscription prices for all newspapers, player or non-player, in the game, and $(r, x, y, D, \xi, \omega, \zeta)$ are analogously defined as vectors of attributes of all newspapers in the game. Variable profit is the difference between revenue and variable cost:

$$\tilde{\pi}^H_j (p, r; x, y, D, \xi, \omega, \zeta; \theta) = \left(p_j q_{j} - \tilde{mc}_{j}^{(q)}q_{j} \right) + \left(r_j a_{j} - \tilde{mc}_{j}^{(a)}a_{j} \right) + \left(\mu_1 q_{j} + \frac{1}{2}\mu_2 q_{j}^2 \right).$$

This is the profit function that is relevant for the decision in the second stage, where publishers observe the product choices, county demographics and the exogenous shocks, $(x, y, D, \xi, \omega, \zeta)$, and choose the optimal prices $(p, r)$. If $\tilde{\pi}^*_j (x, y, D, \xi, \omega, \zeta; \theta)$ and $\tilde{\pi}^j_j (x, y, D, \xi, \omega, \zeta; \theta)$ are equilibrium prices, the overall profit of newspaper $j$ can be expressed as

$$\tilde{\pi}^j_j (x; y, D, \xi, \omega, \zeta, \nu_j; \theta, \tau) = \tilde{\pi}^H_j (\tilde{\nu}^*, \tilde{\nu}; x, y, D, \xi, \omega, \zeta; \theta) - \tilde{fc} (x_j, \nu_j; \tau).$$

### 3.3.4 Necessary Equilibrium Conditions

I now derive the optimality conditions for prices, advertising rates and quality characteristics.\textsuperscript{25} Similar to Rosse (1967), these optimality conditions will be used for identifying the cost structure of newspaper production.

A newspaper publisher has a 2-dimensional price 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 $\tilde{\pi}^H_j$ in (15) with respect to advertising rate $r_j$ yields the optimal display advertising rate as a function of circulation:

$$r_{jt} = \tilde{\zeta} + \frac{\gamma_3}{1 + 1/\lambda_2}q_{jt} + \zeta_{jt}. \quad \text{[RFOC]} (17)$$

Similarly, combining [RFOC](17) and the first order condition with respect to subscription price gives

$$q_m + \frac{\partial q'_m}{\partial p_m} (p_m - mc^{(q)}_m) + \frac{\partial q'_m}{\partial p_m} (\mu_1 + \mu_2 q_m) - \frac{1}{\lambda_2} \frac{\partial a'_m}{\partial p_m} r_m = 0,$$

where $q_m = (q_j, j \in J_m)$ is a vector of circulations of publisher $m$’s newspapers, $(p_m, r_m, a_m, mc^{(q)}_m)$ are analogously defined as the attributes of the newspapers owned by publisher $m$, and $\frac{\partial q'_m}{\partial p_m}$ is the transpose of the Jacobian matrix of $q_m$.\textsuperscript{26} The only difference between (18) and the standard first

\textsuperscript{25}I assume that a pure-strategy Nash equilibrium exists. Finding a set of sufficient conditions for the existence of a Nash equilibrium of this two-stage game is beyond the scope of this paper.

\textsuperscript{26}I follow the standard notation to denote the Jacobian of a function, $g(x) : R^n \rightarrow R^m$, as $\left(\frac{\partial g}{\partial x}\right)_{n \times m}$ to emphasize the correspondence between the columns of the derivative and those in $x'$. 

15
order condition for a multiple product firm involves the last two terms, which are the marginal effect on total advertising profit (preprint profit in the first term and display advertising in the second term) when there is an increase in the newspaper subscription price. Note that $\lambda_2$ is negative.

The first order condition with respect to price (18) holds for all publishers $m \in M$. Inverting $\frac{\partial q'_m}{\partial p_m}$ in (18) gives the estimation equation \textit{(PFOC)}(19):

$$p_{jt} = -\left( \left( \frac{\partial q'_m}{\partial p_m} \right)^{-1} \left( q_m - \frac{1}{\lambda_2} \frac{\partial a'_m}{\partial p_m} r_m \right) \right)_{jt} + \gamma_1 + \gamma_2 f_{jt} + \gamma_3 p_{jt} - \left( \mu_1 + \mu_2 q_{jt} \right) + \omega_{jt}, \forall jt. \quad \text{[PFOC]}(19)$$

When choosing newspaper quality characteristics in the first stage, publishers take into account the impact of their product choice on the equilibrium price in the second stage. The formulation of the optimality condition for the product characteristics therefore requires knowledge of this impact of product choice on the equilibrium price. I take an approach different from the literature\textsuperscript{27} by noticing that knowledge of the gradient of the equilibrium price function at the data points is sufficient to formulate the optimality conditions for the observed product characteristics. Also, the gradient at the observations can be easily computed by taking total derivatives of the first order conditions with respect to the newspaper price and the display advertising rate. Therefore, it is not necessary to compute the equilibrium price for each possible quality choice to obtain the gradient.\textsuperscript{28}

Formally, the necessary optimality condition for the characteristics is that

$$\frac{\partial \sum_{h \in \mathcal{J}_m} \hat{\pi}_h}{\partial x_{kj}} = 0$$

for all newspapers $j \in \mathcal{J}_m$ and all endogenous quality measures $k = 1, \ldots, K$. Each summand is given by

$$\frac{\partial \hat{\pi}_{kj}}{\partial x_{kj}} = \frac{\partial \hat{\pi}_{kj}}{\partial x_{kj}} + \sum_{j' \in \mathcal{J}} \frac{\partial \hat{\pi}_{kj}}{\partial x_{kj}} \frac{\partial p_{j'}}{\partial x_{kj}} + \frac{\partial \hat{\pi}_{kj}}{\partial r_h} \frac{\partial r_h}{\partial x_{kj}} - mc_{kj}(x) (h = j), \forall h, j \in \mathcal{J}_m, \forall k.$$

The first term is the direct impact of increasing characteristic $x_{kj}$ of newspaper $j$ on the variable profit of newspaper $h$ owned by the same publisher. A change in the characteristics of newspaper $j$ also has an impact on the equilibrium subscription prices and advertising rates for all newspapers.

\textsuperscript{27}A common solution in the literature is 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 (2007).

\textsuperscript{28}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 therefore 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 in footnote 27 also requires that the model capture the true profit function accurately so that the equilibrium price function can be accurate.
in a game, which is captured in the second and the third term in the above expression, respectively. Since in the model, the variable profit of newspaper \( h \) \( (\tilde{\pi}_h^{II}) \) does not depend on the advertising rates of other newspapers, the indirect effect of characteristics \( x_{kj} \) on \( \tilde{\pi}_h^{II} \) in the third term is only through affecting the equilibrium advertising rate of newspaper \( h \). This explains the difference between the second and the third term. Finally, the last term is the marginal cost of increasing the characteristic \( x_{kj} \).

In this expression, \( \left( \frac{\partial \tilde{\pi}_h^{II}}{\partial p_j}, \frac{\partial \tilde{\pi}_h^{II}}{\partial r_h}, \frac{\partial \tilde{\pi}_h^{II}}{\partial x_{kj}} \right) \) can be easily computed by taking derivatives of the variable profit function (15). The key is therefore to compute the gradients of the two equilibrium functions \( \frac{\partial \tilde{r}_h^*}{\partial x_{kj}} \) and \( \frac{\partial h^*}{\partial x_{kj}} \). Since the equilibrium price and advertising rate satisfy the first order conditions [RFOC](17) and [PFOC](19), total differentiation of these two equations yields the gradients \( \frac{\partial \tilde{r}_h^*}{\partial x_{kj}} \) and \( \frac{\partial h^*}{\partial x_{kj}} \). See Appendix B for the details. Plugging \( \frac{\partial \tilde{r}_h^*}{\partial x_{kj}} \) and \( \frac{\partial h^*}{\partial x_{kj}} \) into the expression of \( \frac{\partial \tilde{\pi}_h^{II}}{\partial x_{kj}} \) gives the optimality conditions of characteristics, the set of estimation equations [XFOC](20):

\[
\sum_{h \in \mathcal{H}} \left( \frac{\partial \tilde{\pi}_h^{II}}{\partial x_{kj}} \right) + \sum_{j' \in \mathcal{J}_g(jt)} \frac{\partial \tilde{\pi}_h^{II}}{\partial p_{j't}} \frac{\partial \tilde{r}_h^*}{\partial x_{kj}} + \frac{\partial \tilde{\pi}_h^{II}}{\partial x_{kj}} \cdot \frac{\partial \tilde{r}_h^*}{\partial x_{kj}} = \tau_0 + \tau_k x_{kj} + \nu_{kj}, \forall k, jt. \tag{20} \]

### 4 Data

For this study, I have compiled a new data from various sources. See Appendix C for detailed explanation of the data sources and the variable definitions. It 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 \( (a_{jt}, r_{jt}) \).

The data set also contains information on newspaper characteristics. A newspaper is described by the following attributes: (1) news hole \( (x_{1jt}) \), (2) the number of opinion section staff \( (x_{2jt}) \), (3) the number of reporters \( (x_{3jt}) \), (4) frequency of publication \( (f_{jt}) \), and (5) edition (morning or evening newspaper) \( (y_{2jt}) \). Data on all dimensions of the attributes except news hole is available. News hole is the space of a newspaper devoted to news, in other words, pages net of advertising space. As explained in (12), news hole \( x_{1jt} \) can be replaced by \( n_{jt} f_{jt} - \tilde{a} (r_{jt}, q_{jt}, \eta_{jt}; \lambda) \) in the estimation. The latter depends on observable variables and model parameters.

Data on all variables except advertising linage \( (a_{jt}) \), annual subscription price \( (p_{jt}) \) and pages per issue \( (n_{jt}) \) are available for all newspapers in the data period. Display advertising linage data is available for 485 newspaper/years between 1999 and 2005. Therefore, information on this subset
of newspapers is used to identify the advertising demand parameters in [ADV](11). Missing data on price or pages per issue, however, leads to deletion of observations: all newspapers in the game of a newspaper with information on price or pages missing are deleted from the sample. 30 1387 newspaper/years with missing data on price or pages lead to the deletion of 6551 newspaper/years, with 6316 newspaper/years remaining. Summary statistics for the main variables are provided in Table 3 and Table 4.

### Table 3: Summary Statistics of Player Newspapers in Sample

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>median</th>
<th>std</th>
<th>min</th>
<th>max</th>
<th>observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>market penetration (%)</td>
<td>19.13</td>
<td>11.77</td>
<td>18.62</td>
<td>0.3</td>
<td>97.08</td>
<td>2387731</td>
</tr>
<tr>
<td>county distance (100km)</td>
<td>0.71</td>
<td>0.47</td>
<td>0.81</td>
<td>0</td>
<td>6.64</td>
<td></td>
</tr>
<tr>
<td>total circulation</td>
<td>22.729</td>
<td>9.849</td>
<td>43.847</td>
<td>1.132</td>
<td>783,212</td>
<td>631632</td>
</tr>
<tr>
<td>price of newspapers ($)</td>
<td>101.47</td>
<td>97.15</td>
<td>33.75</td>
<td>15</td>
<td>365.31</td>
<td></td>
</tr>
<tr>
<td>price of display advertising ($/column inch)</td>
<td>26.58</td>
<td>13.31</td>
<td>45.19</td>
<td>3.27</td>
<td>748.70</td>
<td></td>
</tr>
<tr>
<td>frequency (issues/52 weeks)</td>
<td>310.70</td>
<td>312</td>
<td>53.87</td>
<td>208</td>
<td>364</td>
<td></td>
</tr>
<tr>
<td>pages (pages/issue)</td>
<td>28.93</td>
<td>23.71</td>
<td>20.79</td>
<td>8</td>
<td>254.57</td>
<td></td>
</tr>
<tr>
<td>opinion staff</td>
<td>2.11</td>
<td>1</td>
<td>2.92</td>
<td>0</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>reporters</td>
<td>22.28</td>
<td>4</td>
<td>43.04</td>
<td>0</td>
<td>218.67</td>
<td></td>
</tr>
</tbody>
</table>

### Table 4: Summary Statistics of the Demographic Characteristics of Counties in Sample

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>median</th>
<th>std</th>
<th>min</th>
<th>max</th>
<th>observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>high education % of pop over 25</td>
<td>17.11</td>
<td>15.22</td>
<td>7.26</td>
<td>5.64</td>
<td>60.48</td>
<td>9357</td>
</tr>
<tr>
<td>median income ($1,000)</td>
<td>34.25</td>
<td>32.85</td>
<td>7.31</td>
<td>16.36</td>
<td>80.12</td>
<td></td>
</tr>
<tr>
<td>median age</td>
<td>36.52</td>
<td>36.70</td>
<td>3.82</td>
<td>20.70</td>
<td>54.30</td>
<td></td>
</tr>
<tr>
<td>urbanization (%)</td>
<td>49.82</td>
<td>50.96</td>
<td>26.51</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Since information on NDM is available only for a few newspapers, I infer NDM from data on county circulations. Specifically, for each newspaper/year $jt$, I sort the counties covered in descending order of county circulation and define NDM as the set of counties that covers 85 percent of total circulation: $C_{jt} = \{(c_1, ..., c_H) \text{ s.t. } \sum_{h=1}^{H} q_{jcht} \geq 0.85q_{jt} \text{ and } \sum_{h=1}^{H-1} q_{jcht} < 0.85q_{jt}\}$, where $q_{jt}$ is the total circulation. 33 With this “definition” of NDM, there are 3994 games in the sample.

---

30 This is because, for example, when the number of pages per issue of newspaper $j$ is not observable, information on its news hole is not available; i.e. its characteristics are unobservable. Hence, for any newspaper $j'$ in $j$’s game, the optimality condition for characteristics conditional on $j$’s opponents choice, including $j$’s choice, is not well-defined. Therefore, $j$’s game are deleted.

31 These observations are at the newspaper/county/year level.

32 These observations are at the newspaper/year level.

33 This is the criterion suggested by the Audit Bureau of Circulation. For the newspapers whose NDM information is observable, this information is consistent with the NDM inferred from county circulation data.
5 Estimation

Five sets of model implications are taken to the data to estimate the model parameters. The model implications derived from newspaper demand [S](8) are used to identify readers’ utility functions and those of advertising demand [ADV](11) are used to identify the dependence of advertising demand on the newspaper’s circulation, its advertising rate and the demographics of its market. Optimality of the observed prices (see [RFOC](17) and [PFOC](19)) is used to identify the variable cost parameters, and optimality of the observed characteristics (see [XFOC](20)) is used to estimate the parameters related to the cost of choosing the characteristics.

The unobservable error terms in the above model implications are solved as functions of the data and the parameters, and then plugged into a set of moment conditions defined by instruments. A Generalized Method of Moments (GMM) estimator is formed based on these moment conditions. The estimation results are presented in Section 6. I now explain the instruments used in the study and how the model parameters are identified.

5.1 Instruments

In the model, newspaper publishers know the unobservable (to econometricians) newspaper-county specific taste \( \xi_{jct} \) and the unobservable cost shocks \((\zeta_{jt}, \omega_{jt}, \nu_j)\) 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 following as instruments for the newspaper quality and the price of newspaper \( j \): the demographics in its own NDM, denoted by \( D_j^{(1)} \), and the demographics in the NDMs of its competitors, denoted by \( D_j^{(2)} \).

To see why \( D_j^{(1)} \) is a valid instrument, first note that consistent with the timing of the model, all unobservable shocks are assumed to be revealed after the NDM of each newspaper is determined and are therefore uncorrelated with \( D_j^{(1)} \). This timing assumption is plausible because location decisions are typically of a longer horizon than both quality and price decisions. But it rules out endogenous NDMs, i.e. endogenous entry/location choices. Secondly, \( D_j^{(1)} \) is correlated with newspaper quality characteristics and prices. This is because county demographics affect the demand for newspapers as well as the advertising demand, which in turn, influence newspaper publishers’ product quality choice and price decisions. Therefore, demographics \( D_j^{(1)} \) can be used as instruments.

However, \( D_j^{(1)} \) is not enough for identification. Even though it does not appear in some estimation equations such as [RFOC](17), it enters other estimation equations such as the mean utility equation [S](8). The demographics in the NDMs of newspaper \( j \)’s competitors, \( D_j^{(2)} \), on the other hand, is excluded from all estimation equations.

To see why \( D_j^{(2)} \) is a valid instrument, first note that for the same reason as \( D_j^{(1)}, D_j^{(2)} \) is
uncorrelated with all unobservables. The intuition for why county demographics in $D^{(2)}_j$ are valid instruments for newspaper prices and qualities is illustrated in Figure 4. The demographics in county 2 ($D_2$) influence the demand for newspaper B as well as its advertising demand, and thus determine the prices and the attributes of this newspaper. Because newspapers A and B are direct competitors, B’s decision on product characteristics and prices affects A’s decision.\(^{34}\) Therefore, the demographics in county 2, $D_2$, 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 a cheap and low-quality newspaper. Thus, the demographics of the NDMs of a newspaper’s competitors are good instruments for this newspaper’s prices and quality characteristics.

### 5.2 Identification

The parameters to be estimated are (i) the parameters in the newspaper demand function, $(\alpha, \beta, \psi, \varphi, \rho, \sigma, \kappa)$; (ii) the parameters in the display advertising demand function $(\phi, \lambda)$; (iii) the cost parameters $(\gamma, \zeta, \tau)$; and (vi) the parameters in the preprint profit function $\mu$.

The identification of $(\alpha, \beta, \psi, \varphi, \sigma)$ in the first set of parameters in [S](8) is similar to that in BLP. In BLP, product characteristics are considered exogenous. They are therefore used as the exogenous source of change in prices and in the choice set facing consumers to identify demand price effects and substitution patterns (which are parameterized by above parameters). In this paper, product characteristics are endogenized. As explained in Section 5.1, I therefore use a different exogenous variation to identify the effects of product characteristics and prices: variation of county demographics.

The parameter that describes the time trend in the outside choice, $\rho$, is identified by the overall change of newspaper circulation over time. Identification of the diminishing utility parameter, $\kappa$, comes from variation of 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

\(^{34}\)This is the intuition behind the instrument used in BLP, where the instruments for the price of A are the characteristics of B, which is assumed to be exogenous there.
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 determined by $\kappa$ as well as the price and quality of the available newspapers.

The second set of parameters is in the advertising demand in [ADV](11). $\phi$ is the vector of the parameters determining the dependence of display advertising on the demographics of a newspaper’s market. It is identified by variation in county demographics. For example, suppose two newspapers have the same circulation and advertising rate, but their circulation areas have different income levels. Any difference in advertising lineage then identifies the parameter $\phi$ corresponding to median income.

$\lambda_1$ and $\lambda_2$ are the display advertising demand elasticities with respect to circulation and advertising rate, respectively. However, $\lambda_1$ and $\lambda_2$ cannot be separately identified from information on advertising lineage only. This is because the source of variation in circulation and advertising is identical. In other words, any exogenous variation that changes circulation also changes the advertising rate. But as the price elasticity of advertising demand, the parameter $\lambda_2$ determines optimal advertising rates for publishers. Therefore, $\lambda_2$ can be identified using the optimality condition with respect to the advertising rate, which then leads to the separate identification of $\lambda_1$ and $\lambda_2$.

It is common in the literature to use firms’ price decisions to identify the cost structure, for example, Rosse (1967). The idea is as follows. With identification of the demand system, the marginal revenue is also identified. Then, the optimal choice of price reveals information on the marginal cost. This is how variable cost parameters ($\gamma, \zeta$) in [PFOC](19) and [RFOC](17) are identified. Similarly, after identification of the variable cost system, the marginal benefit of increasing quality is also identified. Then, the optimal choice of quality characteristics reveals the underlying cost of increasing them, parameterized by $\tau$ in [XFOC](20). For example, suppose an exogenous shock in county demographics or a change in market structure increases the marginal benefit in the variable profit from enlarging the reporter group. Then, the observed change in the number of reporters identifies $\tau_k$ for reporters, i.e. the parameter affecting the marginal cost of increasing reporters.

According to the model, the advertising rate only depends on circulation and the unobservable shocks (to the advertising sales cost) as can be seen in [RFOC](17): $r_{jt} = \tilde{\zeta} + \frac{\gamma_3}{1 + \lambda_2^2/\lambda_2} q_{jt} + \zeta_{jt}$.

Specifically, exogenous variation that leads to changes in circulation identifies $\frac{\gamma_3}{1 + \lambda_2^2/\lambda_2}$ in [RFOC](17), where $\gamma_3$ is the marginal cost of printing one page. $\lambda_2$ and $\gamma_3$ are then separately identified with exogenous variations in county demographics that change circulation but not the number of pages of a newspaper. This can be seen from [PFOC](19). The aforementioned exogenous variations change the marginal effect of increasing price on advertising demand through changing circulation ($\frac{\partial a_m}{\partial p_m}$), but do not change the number of pages printed in a year ($n_{jt} f_{jt}$).
Since the marginal preprint profit can be considered a subsidy to the marginal cost of increasing circulation (see [PFOC](19)), its identification is similar to that of the marginal cost parameters. An exogenous shock to the market size of a newspaper, for example, increases its circulation. Variation in the marginal benefit of increasing circulation, the left hand side of [PFOC](19), which is determined by the identified demand system, then identifies $\mu_2$. The parameter $\mu_1$, however, cannot be separately identified from $\gamma_1$. Recall that marginal cost of increasing circulation is $mc_j(q) = \gamma_1 + \gamma_2 f_j + \gamma_3 n_j f_j + \omega_j$, and the marginal benefit in preprint profit is $\mu_1 + \mu_2 q_j$. As a result, it is $\gamma_1 - \mu_1$ that is relevant for publishers’ decisions, not the values of $\gamma_1$ and $\mu_1$ separately. Hence only the difference can be identified.

6 The Estimation Results

The estimation results are presented in Table 5. The endogenous newspaper characteristic vector, $x_{jt}$, includes news hole, the number of staff for the opinion sections and the number of reporters.\(^{37}\) The estimates of the mean taste ($\beta$) and the disutility from price ($\alpha$) imply that a combination of doubling the news hole of a newspaper and increasing its annual subscription price by 8.5 dollars leaves the mean utility unchanged. Since the estimated reader heterogeneity ($\sigma$) 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 a raise in the subscription price by 140 dollars, and decreasing the number of reporters by half is tantamount to an increase in price by 24.5 dollars. The market size of a newspaper is measured by the logarithm of the number of households in its NDM. The negative sign of $\psi_1$ indicates that readers value a newspaper with for example 10 reporters and covering a small region more than a newspaper that has 10 reporters and serves a large area. County demographics used in this paper include educational level, median income, median age and urbanization,\(^{38}\) of which educational level, age and urbanization positively affect the demand for newspapers. The positive sign of $\rho$ indicates that readers’ utility from subscribing to a newspaper is decreasing over time. This is consistent with the advent of online news, which motivates the inclusion of the time trend in the model.

The parameter $\kappa$ measures the diminishing utility from subscribing to a second newspaper. In a single discrete choice model, this parameter is essentially set to be infinite so that consumers buy at most one product. The estimate of $\kappa$ in this multiple discrete choice model implies that in the majority of the year/county pairs, the percentage of households with two newspapers is close to zero. In the 89 year/county pairs with nontrivial percentage of households with two newspapers, on average 10% of the households with newspapers subscribe to two newspapers.

\(^{37}\)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 $x_{jt}$.

\(^{38}\)See Appendix C for the definitions of these county demographics.
All parameters in the advertising demand function have the expected sign: 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 quality and price of their newspapers after a market structure change.

I set the parameters in the marginal cost of increasing news hole \((\tau_{10}, \tau_{11})\) to zero, because specifications that do not restrict these parameters indicate that news hole does not affect the fixed costs.
cost (fixed with respect to circulation). This can be explained as follows. 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 cost 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.

7 Counterfactual Simulations

In this section, I use counterfactual simulations to study how a change in market structure affects the product choice and price decisions of newspaper publishers. The resulting welfare implications are also investigated. Section 7.1 discusses the welfare measures used: reader surplus, advertiser surplus and publisher surplus. Section 7.2 studies a counterfactual merger of two direct competitors in the Minneapolis market that was blocked by the Department of Justice. This section also analyzes the effect of a merger of two newspapers in this market that do not compete directly, but share a common competitor. A welfare analysis of ownership consolidation in duopoly and triopoly markets, where the publisher of the largest newspaper buys the second largest, is presented in Section 7.3. The correlation of welfare effects and the underlying market structure is also studied. Throughout this section, I use “ownership consolidation” and “merger” interchangeably.

7.1 Welfare Measures

Reader Surplus

The compensation variation for household $i$ in county $c$ in year $t$ is given by

$$CV_{ict} = \frac{V_{ict}^0 - V_{ict}^1}{\alpha},$$

where $\alpha$ is the negative of the household’s marginal value of income, and $V_{ict}^0 - \alpha I_i$ and $V_{ict}^1 - \alpha I_i$ are the expected maximum utility (with respect to the extreme value taste shocks) before and after a merger for household $i$ with income $I_i$. Specifically,$^{39}$

$$V_{ict}^0 = \ln \left( \sum_{j=0}^{J_c} e^{U_{ijct}^0} \right) + \sum_{j=1}^{J_c} \ln \left( \sum_{h \neq 0, j} e^{U_{ihct}^0 - \kappa} + 1 \right) - (J - 1) \ln \left( \sum_{h \neq 0} e^{U_{ihct}^0 - \kappa} + 1 \right),$$

where $U_{ijct}^0 = u_{ijct}^0 - \varepsilon_{ijct}$ is the utility before the merger net of the extreme value taste shock.$^{40}$ $V_{ict}^1$ is analogously defined to $V_{ict}^0$, replacing $U_{ijct}^0$ by $U_{ijct}^1$ and $u_{ijct}^0$ by $u_{ijct}^1$ (utility after merger).

Three welfare measures are reported. (1) Change in the average per-household reader surplus in county $c$ in year $t$ is measured by $\Delta RS_{ct} = E(\xi_i)(CV_{ict})$. (2) Total welfare change is the sum

---

$^{39}$The derivation of this expression follows directly from Small and Rosen (1981) for a single discrete choice model. The only difference is the sum of the second and third term, the expectation (with respect to the extreme value taste shocks) of the second highest utility.

$^{40}$After subtracting the idiosyncratic taste term $\varepsilon_{ijct}$, $U_{ijct}^0$ still depends on the household-specific taste for newspaper characteristics ($\xi_i$), hence the expectation operator.
of the welfare change in all the counties in a game: \[ \Delta RS = \sum_{ct} D_{Lt} \Delta RS_{ct}, \]
where \( D_{Lt} \) is the number of households in county \( c \) in year \( t \). (3) Change in average per-household reader surplus:
\[ \bar{\Delta RS} = \frac{\Delta RS}{\sum_{ct} D_{Lt}}. \]

**Advertiser Surplus**

Since I only observe the advertising lineage for newspapers as a whole, instead of each advertiser’s individual behavior, only the price elasticity of the market demand for advertising is identified, i.e. \( \lambda_2 = \frac{1}{\lambda_3 + \lambda_2 - 1} \). But due to the existence of the negative 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, information on the market demand function is not enough to measure advertiser surplus.

This can be seen as follows. The representative advertiser’s profit function is given by (9) in Section 3.2. Plugging the advertising demand function (10) into the advertiser’s profit function 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 (9)). Since the representative advertiser’s price elasticity parameter \( \lambda_3 \), and the externality parameter, \( \lambda_2 \), cannot be identified separately given only aggregate data, I report the percentage change in advertiser surplus, which is essentially the percentage change in display advertising revenue, \( a_j r_j \).

**Publisher Surplus**

Publisher surplus is given by the profit function in (16).

7.2 Two Case Studies in the Minneapolis/St. Paul Metropolitan Area

**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* and the *St. Paul Pioneer Press*. Three months after the announcement of the transaction, the Department of Justice filed a complaint. Two months later, McClatchy sold the *St. Paul Pioneer Press* to the Hearst Corporation, which later sold it to MediaNews Group. Neither Knight Ridder nor MediaNews owns 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 the *St. Paul Pioneer Press* was simply relabeled.

In this section, I investigate what would have happened to newspaper quality, subscription prices as well as advertising rates and welfare if the ownership consolidation of the *Minneapolis Star Tribune* and the *St. Paul Pioneer Press* had been upheld. These two newspapers are in a game with three other newspapers: the *Faribault Daily News*, the *Stillwater Gazette* and the *St.
Cloud Times. The NDMs of all newspapers are illustrated in Figure 5. The Minneapolis-based Star Tribune and the St. Paul-based Pioneer Press (henceforth, Star and Pioneer) are direct competitors as their NDMs overlap in five counties. Star circulates in a larger area. Whereas the Stillwater Gazette competes with both in Washington County, the Faribault Daily News and the St. Cloud Times compete with Star only.

Figure 5: Newspaper Designated Market (NDM)

(a) NDM of Star Tribune

(b) NDM of Pioneer Press

(c) NDM of St. Cloud Times, Faribault Daily News, Stillwater Gazette
Findings

Table 6 and Table 7 present newspaper quality characteristics, subscription prices and advertising rates at the post-merger equilibrium when only prices are adjusted (Table 6) and when both quality and prices are endogenously chosen by publishers (Table 7).

Table 6: Effect of Ownership Consolidation on Prices and Circulation without Quality Adjustment

<table>
<thead>
<tr>
<th></th>
<th>price ($/year)</th>
<th>ad rate ($/column inch)</th>
<th>circulation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>before</td>
<td>after</td>
<td>change</td>
</tr>
<tr>
<td>Star Tribune</td>
<td>173</td>
<td>182</td>
<td>9</td>
</tr>
<tr>
<td>Pioneer Press</td>
<td>172</td>
<td>204</td>
<td>32</td>
</tr>
<tr>
<td>Stillwater Gazette</td>
<td>78</td>
<td>74</td>
<td>-4</td>
</tr>
<tr>
<td>Faribault Daily News</td>
<td>111</td>
<td>112</td>
<td>1</td>
</tr>
<tr>
<td>St. Cloud Times</td>
<td>150</td>
<td>150</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 7: Effect of Ownership Consolidation on Quality, Prices and Circulation with Quality Adjustment

<table>
<thead>
<tr>
<th></th>
<th>news hole (pages/year)</th>
<th>opinion</th>
<th>reporter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>before</td>
<td>after</td>
<td>change</td>
</tr>
<tr>
<td>Star Tribune</td>
<td>11639</td>
<td>11788</td>
<td>149</td>
</tr>
<tr>
<td>Pioneer Press</td>
<td>12794</td>
<td>14690</td>
<td>1896</td>
</tr>
<tr>
<td>Stillwater Gazette</td>
<td>2325</td>
<td>3620</td>
<td>1295</td>
</tr>
<tr>
<td>Faribault Daily News</td>
<td>7186</td>
<td>7178</td>
<td>-8</td>
</tr>
<tr>
<td>St. Cloud Times</td>
<td>14759</td>
<td>14511</td>
<td>-248</td>
</tr>
<tr>
<td></td>
<td>price ($/year)</td>
<td>ad rate ($/column inch)</td>
<td>circulation</td>
</tr>
<tr>
<td>Star Tribune</td>
<td>173</td>
<td>181</td>
<td>8</td>
</tr>
<tr>
<td>Pioneer Press</td>
<td>172</td>
<td>198</td>
<td>26</td>
</tr>
<tr>
<td>Stillwater Gazette</td>
<td>78</td>
<td>40</td>
<td>-38</td>
</tr>
<tr>
<td>Faribault Daily News</td>
<td>111</td>
<td>112</td>
<td>1</td>
</tr>
<tr>
<td>St. Cloud Times</td>
<td>150</td>
<td>150</td>
<td>0</td>
</tr>
</tbody>
</table>

From the table, we can see that: (1) Without quality adjustment, both Star and Pioneer increase their subscription prices. (2) With quality adjustment, both parties of the merger increase news hole, reduce the number of opinion-section staff and reporters, and decrease the overall newspaper quality.41 (3) The adjustment of the smaller party of the merger (Pioneer) is much bigger than that of the larger party (Star) in both scenarios — with or without quality adjustment. (4) In both scenarios, the Stillwater Gazette, which competes with the two parties of the merger, lowers its subscription price. Its quality increases when quality adjustment is allowed. (5) The magnitude of the price adjustment in the two scenarios are different: the price adjustment for the two parties of the merger is smaller when quality adjustment is allowed, while that for the Stillwater Gazette is larger. In other words, ignoring quality adjustment leads to an overestimation of the price ad-

41Since the estimated reader heterogeneity is small, overall quality of newspaper j can be defined by the mean utility from each characteristic: $\beta_1 \log(1 + x_{1j}) + \beta_2 \log(1 + x_{2j}) + \beta_3 \log(1 + x_{3j})$. 

27
justment for the newspapers involved in the merger and an underestimation for the newspaper competing with the merged newspapers. (6) In both scenarios, the Faribault Daily News and the St. Cloud Times only adjust marginally after the merger.

Intuition Underlying the Findings

A detailed explanation of the economic interactions that drive these results can be found in Appendix D. Here, I provide some intuition underlying these observations.

(1) After the publisher of Star, McClatchy, purchases Pioneer, it internalizes the positive price cross-effect of these two newspapers: a higher price of Star, for example, leads to an increase in the market share of Pioneer and therefore raises its profit. This explains observation (1).

(2) Analogously, there also exist quality cross-effects. As will be shown in Appendix D, the cross-effect of news hole is positive and the cross-effects of opinion staff and reporters are negative at the pre-merger equilibrium. The sign of the cross-effect of news hole can be positive\(^{42}\) because different from the other two characteristics, news hole also affects the marginal cost of increasing circulation. Specifically, increasing news hole leads to a higher marginal cost \(mc^{(q)}\) and hence not always to an advantage over other newspapers in the price competition.\(^{43}\) This is consistent with observation (2).

(3) As will be explained in Appendix D, the 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 the circulation from its small newspapers to large newspapers. Here, even though McClatchy decreases the quality of both newspapers, it adjusts the smaller newspaper by a bigger margin due to this incentive.

(4) An increase in the prices or a decrease in the quality of Star and Pioneer, the competitors of the Stillwater Gazette, leads to an increase in the latter’s marginal benefit from increasing circulation. It thus raises its circulation by decreasing its price and increasing its quality as in observation (4).

(5) As the publisher McClatchy internalizes the overall negative quality cross-effect between Star and Pioneer, it decreases the quality of the two newspapers involved in the merger. Also the positive price cross-effect is weakened. Therefore, price adjustments are smaller when quality adjustment is allowed. Similarly, as the quality of Star and Pioneer decreases, the marginal benefit for the Stillwater Gazette from decreasing its subscription price is even higher. Therefore, its price adjustment is larger when quality adjustment is allowed.

\(^{42}\) Appendix D shows that the sign of the cross-effect of news hole is not determinate. In this merger, it is positive at the original equilibrium.

\(^{43}\) Quality cross-effects here take into account the impact of quality on the equilibrium price.
Finally, observation (6) is explained by the NDMs of the newspapers involved (see Figure 5). The *Faribault Daily News* increases its price marginally because it only competes with *Star*, which does not change much after the merger. Similarly, because *Star* does not have a strong presence in the NDM of the *St. Cloud Times*, the *St. Cloud Times* almost does not adjust its price either.

**Welfare Implications and Comparison to a Merger without Quality Adjustment**

These adjustments in quality and subscription price influence the circulation of each newspaper and hence the optimal advertising rate. All these changes decrease the overall reader surplus by 7.94 million dollars, advertiser surplus by 5.59% and increase publisher surplus by 0.52 million dollars. So, the total welfare declines. Specifically, as Table 8 shows, households in all counties except Sterns County are worse-off. There the dominating newspaper *St. Cloud Times* increases its quality slightly. Across all counties, the average per-household reader surplus (\(\Delta RS\)) declines by 6 dollars. Counties covered by the two merged parties are affected the worst. For example, readers’ welfare falls by 15 dollars in Ramsey County, which is the home county of *Pioneer* and close to Hennepin County, the home county of *Star*. The profits of all publishers increase. The profit of the originally smallest newspaper, the *Stillwater Gazette*, increases by a larger margin than the two median-sized newspapers because the latter interact only marginally with the two parties of the merger. *Star* and *Pioneer* lower the quality and leave space for the *Stillwater Gazette* to increase quality. In fact, it even overtakes the *Faribault Daily News* in terms of circulation.

<table>
<thead>
<tr>
<th>(a) Welfare Implications</th>
<th>(b) Change in Publisher Surplus</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\Delta RS)</td>
<td>-7.94 million</td>
</tr>
<tr>
<td>(%\Delta AS)</td>
<td>-5.59%</td>
</tr>
<tr>
<td>(\Delta PS)</td>
<td>0.52 million</td>
</tr>
<tr>
<td>newspapers</td>
<td>(\Delta PS)</td>
</tr>
<tr>
<td>Star &amp; Pioneer</td>
<td>374000</td>
</tr>
<tr>
<td>Stillwater Gazette</td>
<td>84460</td>
</tr>
<tr>
<td>Faribault Daily News</td>
<td>29500</td>
</tr>
<tr>
<td>St. Cloud Times</td>
<td>24110</td>
</tr>
</tbody>
</table>

(c) Average Change in Reader Surplus per Household

<table>
<thead>
<tr>
<th>county</th>
<th>(\Delta RS_{ct})</th>
<th>county</th>
<th>(\Delta RS_{ct})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anoka</td>
<td>-1.36</td>
<td>Rice</td>
<td>-3.18</td>
</tr>
<tr>
<td>Benton</td>
<td>-0.70</td>
<td>Scott</td>
<td>-3.74</td>
</tr>
<tr>
<td>Carver</td>
<td>-3.25</td>
<td>Sherburne</td>
<td>-1.59</td>
</tr>
<tr>
<td>Dakota</td>
<td>-9.83</td>
<td>Stearns</td>
<td>0.43</td>
</tr>
<tr>
<td>Hennepin</td>
<td>-4.48</td>
<td>Washington</td>
<td>-5.44</td>
</tr>
<tr>
<td>McLeod</td>
<td>-2.02</td>
<td>Wright</td>
<td>-2.30</td>
</tr>
<tr>
<td>Ramsey</td>
<td>-14.58</td>
<td>St. Croix, WI</td>
<td>-9.10</td>
</tr>
</tbody>
</table>

The welfare change without quality adjustment is -7.93 million for readers, -4.96% for advertisers and 0.91 million for publishers. Therefore, ignoring quality adjustment overestimates the gain in
publisher surplus, and underestimates the loss in reader and advertiser surplus in this merger. Note that an overestimation of the price adjustment (see observation (5)) can be consistent with an underestimation of the loss in readers’ welfare. Even though the price adjustment is smaller with quality adjustment, the quality of the newspapers is lower as well. It is the overall utility from the newspapers that determines readers’ welfare. Section 7.3 analyzes the relationship between the bias in estimating the welfare effect from ignoring quality adjustment and the underlying market structure. In particular, I show that the bias in the estimate for the change in reader surplus can be significantly larger than 10,000 dollars.

**Case 2. Ownership Consolidation of Indirect Competitors**

In the above ownership consolidation study, the two parties of the merger are direct competitors. This is usually the main focus in policy analyses. In fact, a similar quality and price cross-effect exists even when the merged parties just share a common competitor. Therefore, an ownership consolidation of such two newspapers also affects the quality and prices of the newspapers involved. To illustrate this point and quantify the effect, I study a counterfactual merger of *Pioneer* and the *St. Cloud Times* who do not compete directly as shown in the NDMs of these two newspapers in Figures 5(b) and 5(c).

Table 9: The Effect of the Ownership Consolidation of *Pioneer* and *St. Cloud Times* on Quality, Prices and Circulations with Quality Adjustment

<table>
<thead>
<tr>
<th></th>
<th>news hole (pages/year)</th>
<th>opinion</th>
<th>reporter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>before</td>
<td>after</td>
<td>change</td>
</tr>
<tr>
<td>Star Tribune</td>
<td>11639</td>
<td>11638</td>
<td>-1</td>
</tr>
<tr>
<td>Pioneer Press</td>
<td>12794</td>
<td>12802</td>
<td>8</td>
</tr>
<tr>
<td>Stillwater Gazette</td>
<td>2325</td>
<td>2327</td>
<td>2</td>
</tr>
<tr>
<td>Faribault Daily News</td>
<td>7186</td>
<td>7186</td>
<td>0</td>
</tr>
<tr>
<td>St. Cloud Times</td>
<td>14759</td>
<td>15512</td>
<td>753</td>
</tr>
<tr>
<td>price ($/year)</td>
<td>173</td>
<td>173</td>
<td>0</td>
</tr>
<tr>
<td>Pioneer Press</td>
<td>172</td>
<td>171</td>
<td>-1</td>
</tr>
<tr>
<td>Stillwater Gazette</td>
<td>78</td>
<td>78</td>
<td>0</td>
</tr>
<tr>
<td>Faribault Daily News</td>
<td>111</td>
<td>111</td>
<td>0</td>
</tr>
<tr>
<td>St. Cloud Times</td>
<td>150</td>
<td>151</td>
<td>1</td>
</tr>
</tbody>
</table>

The results are presented in Table 9. Again, the smaller party of the merger adjusts more than the larger party. As the *St. Cloud Times* increases news hole by about 2 pages per issue and reduces the number of opinion-section staff and reporters, its overall quality falls. Therefore, households in the counties served by it are worse off. The impact of such an ownership consolidation is much smaller than that of merging two direct competitors. The loss in readers’ welfare, for example, is 18 cents on average and 3 dollars in the county that is worst affected. Overall, reader surplus
decreases by 0.22 million, publisher surplus increases by 0.02 million and the change in advertiser surplus is negligible.

7.3 Welfare Analysis of Duopoly Mergers and Triopoly Mergers

In this section, I study welfare implications of ownership consolidations in duopoly and triopoly markets. In a duopoly merger, the publisher of one newspaper buys the other and becomes a monopoly in the market. A triopoly merger is defined as the publisher of the largest newspaper buying the second largest. The welfare effect of an ownership consolidation in a market depends on the details of the market structure. I will present the distribution of the welfare effects for all duopoly and triopoly markets in the 2005 sample, and then examine how they vary with market characteristics.

Figures 6 and 7 show welfare changes after an ownership consolidation in 40 duopoly markets and 13 triopoly markets in the 2005 sample, the last year in the data set. The markets are sorted according to $\Delta RS$, the change in average per-household reader surplus with quality adjustment. A dot in the upper-left graph of Figure 6, for example, represents $\Delta RS$ in a market when quality adjustment is allowed. A cross represents $\Delta RS$ without quality adjustment. The difference between a cross and a dot on the same vertical line therefore represents the bias in estimating $\Delta RS$ when quality adjustment is ignored. The upper-right graph plots changes in total reader surplus ($\Delta RS$). For example, in the market shown in the middle of the graph, the total reader surplus drops by more than 5 million dollars after the merger when quality adjustment is allowed, and by around 2 million dollars without quality adjustment. Ignoring quality adjustment therefore underestimates the total readers’ welfare loss by around 3 million dollars. The lower-left graph and the lower-right graph show percentage changes in advertiser surplus and changes in publisher surplus in millions, respectively. Finally, Figure 7 does the same for the 13 triopoly markets.

The median changes in different welfare measures are presented in Table 10. The total welfare falls unambiguously in 38 duopoly markets. The total welfare in all triopoly markets simulated falls after the merger. However, the presence of a competitor mitigates the welfare loss for readers and advertisers. This is because the merged parties downgrade the quality of their newspapers by a smaller margin than they would have done without a competitor and some even increase newspaper quality (in 6 markets). This mitigation in welfare loss is also partially due to the competitors sometimes increasing the quality. Figures 6 and 7 show that ignoring quality adjustment typically leads to an underestimation of the loss in reader surplus and the gain in publisher surplus. The median bias in estimating welfare is 4 dollars per household in triopoly mergers and 2 dollars in duopoly mergers.

---

44Reader surplus falls in all duopoly markets. As expected, publisher surplus increases uniformly. But the net change in the sum of reader surplus and publisher surplus is negative in all 40 duopoly markets simulated. Among 11 markets where the circulation of at least one newspaper increases, 2 markets witness an increase in advertiser surplus...
Figure 6: Welfare Implications of Duopoly Mergers

Figure 7: Welfare Implications of Triopoly Mergers
Table 10: Median Welfare Changes across Duopoly and Triopoly Markets

<table>
<thead>
<tr>
<th></th>
<th>$\Delta RS$ (dollars)</th>
<th>$\Delta RS$ (millions)</th>
<th>%$\Delta AS$ (%)</th>
<th>$\Delta PS$ (millions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duopoly with quality adjustment</td>
<td>-16.32</td>
<td>-0.97</td>
<td>-6.12%</td>
<td>0.10</td>
</tr>
<tr>
<td>Duopoly without quality adjustment</td>
<td>-10.52</td>
<td>-0.64</td>
<td>-3.04%</td>
<td>0.07</td>
</tr>
<tr>
<td>Triopoly with quality adjustment</td>
<td>-5.11</td>
<td>-0.83</td>
<td>-3.06%</td>
<td>0.09</td>
</tr>
<tr>
<td>Triopoly without quality adjustment</td>
<td>-5.82</td>
<td>-0.61</td>
<td>-1.85%</td>
<td>0.08</td>
</tr>
</tbody>
</table>

To examining these welfare changes more closely, I now study the following: (1) variation of the change in average per-household reader surplus ($\Delta RS$) across markets; (2) variation of the bias in $\Delta RS$ when quality adjustment is ignored ($\Delta \Delta RS = (\Delta RS$ without quality adjustment) - ($\Delta RS$ with quality adjustment)); and (3) the difference between duopoly and triopoly markets. To understand how welfare effects vary across markets, I run two regressions of $\Delta RS_m$ and $\Delta \Delta RS_m$, where the subscript $m$ represents a market, on some market characteristic variables. As explained before, market structure cannot be captured by simple indices. I therefore regress $\Delta RS_m$ and $\Delta \Delta RS_m$ on a triopoly dummy and endogenous variables, which are correlated with the underlying market characteristics that determine the welfare change. The regression therefore shows a correlation pattern rather than a causal effect.

The result of the first regression is presented below. Standard errors are in parentheses.

$$
\Delta RS_m = 25.44 - 0.99 \text{pen}_m - 0.28 \text{overlap}_m^{(1)} + 4.67 \log \left( \frac{q_{1m}}{q_{2m}} \right) + 5.04 tr_i_m + 0.25 tr_i_m \cdot \text{overlap}_m^{(2)} + \varphi_m
$$

(7.77) (0.16) (0.05) (2.04) (3.74) (0.12)

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 do not affect their welfare much. The pre-merger newspaper penetration rate ($\text{pen}_m$), measured by the ratio of the total newspaper circulation and the number of households in market $m$, is used to capture this aspect of the market. The negative sign in the estimation is as expected: readers’ welfare loss ($-\Delta RS_m$) increases when readers care about newspapers. An increase in the penetration rate by 1 percentage point leads to an increase in the average welfare loss per household of 99 cents.

Another market feature that affects $\Delta RS_m$ is the importance of the two merged parties’ common circulation area to these two newspapers. This influences how strong the cross-effect is. Suppose two newspapers only compete with each other in a county that is far away from their home counties. Then, this county might not play a large role in generating profit for these two newspapers because

\[pen_m\] also captures that for a given change in quality and prices of newspapers, the welfare change is decreasing in pre-merger utilities, i.e. the welfare loss is increasing in pre-merger utilities. This can be seen from the welfare measure in Section 7.1.
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 largest newspapers in the market: 

\[ \text{overlap}^{(1)}_m = 100 \times \left( \sum_{c \in \text{CTY}_{1,2}} q_{1mc} / q_{1m} \right), \]

where \( \text{CTY}_{1,2} \) is the intersection of the NDMs of the two largest newspapers, and \( q_{1mc} \) and \( q_{1m} \) are county circulation (in county \( c \)) and total circulation of the largest newspaper in the market, respectively.\(^{46}\)

The above regression indicates a negative correlation between \( \Delta \Delta R S_m \) and \( \text{overlap}^{(1)}_m \), meaning that the larger the overlapping is, the larger is the welfare loss for readers.

The third factor is the pre-merger asymmetry of the two parties of the merger in terms of circulation, measured by \( \log \left( q_{1m} / q_{2m} \right) \). As explained in Section 7.2, due to increasing marginal benefit of a higher circulation, the publisher of the merged parties will not adjust the quality and prices of the larger party by much. Since an adjustment in a larger newspaper has a bigger impact on readers’ welfare than the same adjustment in a smaller newspaper, asymmetry matters, and specifically, the larger the asymmetry is, the smaller is the welfare loss for readers, as indicated by the positive sign in the above regression.

Finally, as explained before, the presence of a competitor mitigates the welfare loss for readers and advertisers, because the merged parties decrease the quality of their newspapers by a smaller margin when facing a competitor. Therefore, the triopoly dummy has a positive sign in the regression. How strong the competition effect is depends on how strong the cross-effect between the two merged parties and their competitor, which is again captured by the pre-merger overlapping rate:

\[ \text{overlap}^{(2)}_m = 100 \times \left( \sum_{c \in \text{CTY}_{1,3}} q_{1mc} + \sum_{c \in \text{CTY}_{2,3}} q_{2mc} \right) / (q_{1m} + q_{2m}), \]

where \( \text{CTY}_{1,3} \) is the intersection of the NDMs of the largest newspaper and the competitor, and \( \text{CTY}_{2,3} \) is analogously defined for the second largest newspaper and the competitor. The bigger the overlapping, the stronger the competition effect, and thus the smaller the welfare loss.

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

\[
\Delta \Delta R S_m = 2.71 - 4.86 \text{tri}_m + 0.30 \text{pen}_m - 0.23 s_{1m} + \varphi_m
\]

(5.23) (2.18) (0.11) (0.11)

Again, triopoly dummy and penetration rate matter. For example, a positive sign in front of \( \text{pen}_m \) means the higher the penetration rate, the larger the bias in measuring welfare loss, in other words, the bigger the problem of ignoring quality adjustment. Another factor that determines \( \Delta \Delta R S_m \) is the demand elasticity with respect to price. To understand this, denote the post-merger/with-quality-adjustment equilibrium by \( (p^1, x^1) \) and that without quality adjustment by \( (p^2, x^0) \), where \( x^0 \) is a vector of the pre-merger quality of all newspapers in the market. Given

\(^{46}\)The pre-merger overlapping rate can be also defined for the second largest newspaper in the market as \( 100 \times \left( \sum_{c \in \text{CTY}_{1,2}} q_{2mc} / q_{2m} \right) \), 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.
that the estimated reader heterogeneity is small, what matters for readers’ welfare is the mean utility component \( p_j \alpha + x_j \beta \). I now explain how demand elasticity with respect to price affects \( (p_j^2 \alpha + x_j^0 \beta) - (p_j^1 \alpha + x_j^1 \beta) \). When a publisher is prevented from setting its at \( x_j^1 \) and has to stay at \( x_j^0 \), it can increase price by \( (x_j^0 - x_j^1) \beta / (-\alpha) \), while keeping the mean utility and thus its circulation unchanged. But the publisher’s goal is to maximize its profit instead of keeping the circulation at a certain level. It will therefore continue to increase the price, and thus decrease the mean utility, until the marginal profit from doing so is 0. How much it will increase price beyond \( (x_j^0 - x_j^1) \beta / (-\alpha) \), i.e. the difference between \( (p_j^2 - p_j^1) \) and \( (x_j^0 - x_j^1) \beta / (-\alpha) \), depends on the price elasticity of demand. Therefore, how much readers’ welfare will be affected depends on the price elasticity as well. A large elasticity leads to a small increase in price and hence a small decrease in the mean utility of the newspaper. The welfare effect of a change from \( (p^1, x^1) \) to \( (p^2, x^0) \) for readers is therefore small. Since the price elasticity in logit models depends positively on market shares when market shares are smaller than 1/2, I use \( s_{1m} \), the pre-merger market penetration rate of the largest newspaper in its largest circulation county, to capture this factor. The sign in the regression result is consistent with the conjecture: a higher price elasticity of demand leads to smaller welfare changes induced by a change from \( (p^1, x^1) \) to \( (p^2, x^0) \), i.e. a smaller bias from ignoring quality adjustment.

8 Conclusion

In this paper, I set up a structural model and use counterfactual simulations to study the welfare implications of 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 a direct and an indirect merger in the Minneapolis market. I also quantify the welfare implications of ownership consolidations 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 decreases the overall quality and increases the prices of both newspapers. The adjustment of the St. Paul Pioneer Press, the smaller newspaper, is much larger than that of the bigger newspaper because advertising profit is convex in circulation and thus a multiple-newspaper publisher has an incentive to shift circulations to its larger newspaper.

Second, the simulation results show that the median loss in reader surplus in duopoly mergers is 16 dollars per household and 5 dollars in triopoly mergers. Readers’ welfare loss of 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 existence of a competitor mitigates the loss in readers’ welfare due to a competition effect. The larger the competition effect is, the smaller the welfare loss.

Third, ignoring quality adjustment typically leads to an underestimation of the loss in reader surplus and the gain in publisher surplus. In general, the bias in measuring the welfare effect of ownership consolidations is smaller in a triopoly merger and when the price elasticity of newspaper demand is higher. It is larger when households care more about reading newspapers.

Fourth, ownership consolidation has an impact on quality choice and thus welfare even when the newspapers involved in the merger do not compete directly. This welfare effect, however, is more than an order of magnitude smaller than in the ownership consolidation of direct competitors in the simulated market.
References


Dirks, Van Essen & Murray (2005), “Number of independent dailies steadily declines.”


Appendix

A Proof of Theorem 1

In this appendix, I show that the invertibility result in BLP can be extended to a multiple discrete choice model. I only prove Theorem 1 for a multiple discrete choice model where the number of discrete choices is limited to at most two. An extension of the result to a model where consumers can choose up to \( \bar{n} \leq J \) products, where \( J \) is the total number of products available in a choice set, is available upon request.

The proof is similar to that in BLP, where a key step is to show the existence of an upper bound for the fixed point of the mapping \( F \). The main difference between the proof here and the BLP proof is at this step. In a multiple discrete choice model, the value of \( \delta_j \) that solves \( \sum_{h=1}^J s_h = \sum_{h=1}^J \tilde{s}_h (\delta, \mathbf{x}; P_\zeta, \sigma, \kappa) \) where \( \delta_k = -\infty \) for \( \forall k \neq j \) is no longer necessarily the upper bound of \( \delta_j \). This value does not even exist when the left hand side \( \sum_{h=1}^J s_h > 1 \).\footnote{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.}

Note that

\[
\sup_{\delta=(\delta_1,\ldots,\delta_J), \delta_k=-\infty \text{ for } k \neq j} \sum_{h=1}^J \tilde{s}_h (\delta, \mathbf{x}; P_\zeta, \sigma, \kappa) = 1.
\]

The theorem is proved in two steps. All statements below are true for any given \((\mathbf{x}, P_\zeta, \sigma, \kappa)\). Therefore, these arguments in \( \tilde{s}_j \) are omitted for presentational simplicity.

Claim 1 There exist \( \underline{\delta} \) and \( \bar{\delta} \) such that if \( F \) has a fixed point \( \delta^* \), \( \delta^* \) must be in \([\underline{\delta}, \bar{\delta}]^J\).

Proof. Construction of the lower bound \( \underline{\delta} \) is the same as in BLP. As will be shown in the proof of Claim 2, \( F_j (\delta) \) is increasing in all dimensions of \( \delta \). Define \( \delta_j = \lim_{\delta \to -\infty} F_j (\delta) = \int \exp (\mathbf{x}; \zeta) dP_\zeta (\zeta; \sigma) \) if \( \delta^* \) is a fixed point of \( F \), \( \delta^*_j = F_j (\delta^*) \geq \delta^*_j = \min_{j'} (\delta^*_{j'}) \).

Note that \( \delta^* \) as a fixed point of \( F \) satisfies \( \sum_{j=1}^J \tilde{s}_j (\delta^*) = \sum_{j=1}^J s_j \). If two or more dimensions of \( \delta \) go to infinity, \( \sum_{j=1}^J \tilde{s}_j (\delta) \) approaches 2. But \( \sum_{j=1}^J s_j < 2 \). So, there exists at most one \( j \) such that \( \delta^*_j \) is unbounded. Given that all other \( \delta_k, k \neq j \) are bounded, \( \lim_{\delta \to -\infty} \tilde{s}_j (\delta) = 1 \). So, \( \delta^*_j \) has to be bounded as well to ensure \( \delta^*_j = s_j < 1 \). Therefore, all dimensions of \( \delta^* \) are bounded. Let the upper bound be \( \delta' \). Define \( \bar{\delta} = \delta' + 1 \).

Claim 2 \( F : [\underline{\delta}, \bar{\delta}]^J \to R^J \) has a unique fixed point.

Proof. Define \( \tilde{F} : [\underline{\tilde{\delta}}, \bar{\tilde{\delta}}]^J \to R^J \) as \( \tilde{F} (\delta) = \min (F(\delta), \bar{\delta}) \). Since \( \underline{\delta} \) is the lower bound of \( F_j (\delta) \), \( \tilde{F} \) is a mapping from \([\underline{\tilde{\delta}}, \bar{\tilde{\delta}}]^J \) to itself. According to BLP, if (i) \( \partial F_j (\delta) / \partial \delta_h \geq 0 \) for any \( h, j \) and (ii) \( \sum_{h=1}^J \partial F_j (\delta) / \partial \delta_h < 1 \) for any \( j \), then \( \tilde{F} \) is a contraction mapping.

To show that these two conditions hold, first note that \( 0 < \tilde{\Psi}_j (1), \tilde{\Psi}_j (2), \tilde{\Psi}_j (3) \) \footnote{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.} and \( \tilde{\Psi}_j (3) < 1 \). When \( \delta \) is in a bounded set, \( \tilde{\Psi}_j (1) \geq \tilde{\Psi}_j (3) \) for all \( j \) and \( \tilde{\Psi}_j (2) \geq \tilde{\Psi}_j (3) \) for \( j = j' \neq j \).

(i) I first show that condition (i) holds when \( h = j \). Note that

\[
\frac{\partial \tilde{s}_j}{\partial \delta_j} = \int \tilde{\Psi}_j (1 - \tilde{\Psi}_j (1)) dP_\zeta (\zeta; \sigma) + \sum_{j' \neq j} \int [\tilde{\Psi}_j (2 (1 - \tilde{\Psi}_j (1) - \tilde{\Psi}_j (2)) (1 - \tilde{\Psi}_j (3))] dP_\zeta (\zeta; \sigma) .
\]
In the first summand, \( \tilde{\Psi}_j^{(1)}(1 - \tilde{\Psi}_j^{(1)}) < \tilde{\Psi}_j^{(1)} \). In the second summand, 
\[
\left[ \tilde{\Psi}_{j,j'}^{(2)}(1 - \tilde{\Psi}_{j,j'}^{(2)}) - \tilde{\Psi}_j^{(3)}(1 - \tilde{\Psi}_j^{(3)}) \right] \leq \left[ \tilde{\Psi}_{j,j'}^{(2)} - \tilde{\Psi}_j^{(3)} \right].
\]
Therefore, \( \partial \tilde{s}_j/\partial \delta_j < \tilde{s}_j \) and \( \partial F_j(\delta) / \partial \delta_j = 1 - (\partial \tilde{s}_j/\partial \delta_j) / \tilde{s}_j > 0. \)

For \( h \neq j \),
\[
\begin{align*}
\frac{\partial \tilde{s}_j}{\partial \delta_h} & = -\int \tilde{\Psi}_j^{(1)} \rho_h \ dP_\zeta (\zeta; \sigma) + \int \tilde{\Psi}_j^{(3)} \rho_h \ dP_\zeta (\zeta; \sigma) \\
& + \sum_{j' \neq j,h} \int \left( -\tilde{\Psi}_{j,j'}^{(2)} \tilde{\Psi}_{h,j'}^{(2)} + \tilde{\Psi}_j^{(3)} \tilde{\Psi}_h^{(3)} \right) dP_\zeta (\zeta; \sigma) \\
& \leq \sum_{j' \neq j,h} \int \left( -\tilde{\Psi}_{j,j'}^{(2)} \tilde{\Psi}_{h,j'}^{(2)} + \tilde{\Psi}_j^{(3)} \tilde{\Psi}_h^{(3)} \right) dP_\zeta (\zeta; \sigma)
\end{align*}
\]
Therefore, \( \partial F_j(\delta) / \partial \delta_h = - (\partial \tilde{s}_j/\partial \delta_h) / \tilde{s}_j \geq 0. \)

(ii) \( \sum_{h=1}^{J} \partial F_j(\delta) / \partial \delta_h = 1 - \sum_{h=1}^{J} (\partial \tilde{s}_j/\partial \delta_h) / \tilde{s}_j \) and \( \sum_{h=1}^{J} \frac{\partial \tilde{s}_j(\delta)}{\partial \delta_h} = \frac{\partial \tilde{s}_j(\delta + \Delta)}{\partial \Delta} \big|_{\Delta = 0}. \)
\[
\begin{align*}
\frac{\partial \tilde{s}_j(\delta + \Delta)}{\partial \Delta} \big|_{\Delta = 0} & = \int \left( \tilde{\Psi}_j^{(1)} \right)^2 \frac{1}{e^{\delta_j + \Delta \zeta}} dP_\zeta (\zeta; \sigma) \\
& + \sum_{j' \neq j} \int \left[ \left( \tilde{\Psi}_{j,j'}^{(2)} \right)^2 - \left( \tilde{\Psi}_j^{(3)} \right)^2 \right] \frac{e^{\delta_j + \Delta \zeta}}{e^{\delta_j + \Delta \zeta}} dP_\zeta (\zeta; \sigma) \\
& > 0.
\end{align*}
\]
Therefore, \( \sum_{h=1}^{J} \partial F_j(\delta) / \partial \delta_h < 1. \)

According to BLP, (i) and (ii) implies that \( \tilde{F} \) is a contraction mapping from \( [\tilde{\delta}, \tilde{\delta}]^J \) to \( [\tilde{\delta}, \tilde{\delta}]^J \). Note that \( ([\tilde{\delta}, \tilde{\delta}]^J, \| \cdot \|) \) is a complete metric space. The contraction mapping \( \tilde{F} \) therefore has a unique fixed point. Denote it by \( \delta^* \). Hence, \( \tilde{F}(\delta^*) = \min \left( F(\delta^*), \tilde{\delta} \right) = \delta^* \). Claim 1 shows that if \( \delta_j^* = \tilde{\delta} \) for any \( j \), there exists \( k \) such that \( F_k(\delta^*) \neq \delta^* \). Moreover, the proof of Claim 1 implies that \( F_k(\delta^*) < \delta^*_k \). Therefore, \( F_k(\delta^*) = F_k(\delta^*) < \delta^*_k \), which is a contradiction to \( \delta^* \) being a fixed point of \( \tilde{F} \). So, the unique fixed point of \( \tilde{F} \) cannot be on the bound, and hence it is also the unique fixed point of \( F(\delta) \) on \( [\tilde{\delta}, \tilde{\delta}]^J \). □

Claim 1 implies that the unique fixed point in claim 2 is also the unique fixed point of \( F : \mathbb{R}^J \rightarrow \mathbb{R}^J \). This completes the proof of the theorem.
B  Jacobian of the Equilibrium Price Functions

The jacobian of the equilibrium price function is obtained by total differentiation of the two optimal pricing conditions. To see the details, combine the optimality conditions (18) for all \( m \in \mathcal{M} \) as follows.

\[
\tilde{q} + \left( \frac{\partial \tilde{q}}{\partial p} \right)^\# \left( p - \overline{mc}^{(q)} + \mu_1 + \mu_2 q \right) - \frac{1}{\lambda_2} \left( \frac{\partial \tilde{q}}{\partial p} \right)^\# r = 0,
\]

where the \((j,k)\)-element of \( \left( \frac{\partial \tilde{q}}{\partial p} \right)^\# \) is \( \frac{\partial \tilde{q}_{jk}}{\partial p} \), if newspaper \( j \) and \( k \) are owned by the same newspaper publisher and 0 otherwise. In other words, notation \((X)^\#\) represents the element-wise multiple of matrix \( X \) and a dummy matrix defined by ownership. When the optimal display advertising rate \( \hat{r}_{kj} (\hat{q}_j, \zeta_j) \) defined in [RFOC](17) is plugged in, the above equation can be rewritten as

\[
\tilde{m} (\delta, y, x, \xi, \zeta, \omega) = \tilde{m}_1 (\delta, x, y, p, \xi) + \tilde{m}_2 (\delta, y, x, \xi, \zeta, \omega) = 0
\]

where

\[
\tilde{m}_1 (\delta, x, y, p, \xi) = \left( \frac{\partial \tilde{q}}{\partial p} \right)^\# p
\]

and

\[
\tilde{m}_2 (\delta, y, x, \xi, \zeta, \omega) = \tilde{q} - \left( \frac{\partial \tilde{q}}{\partial p} \right)^\# \left( \overline{mc}^{(q)} - \mu_1 - \mu_2 q \right) - \frac{1}{\lambda_2} \left( \frac{\partial \tilde{q}}{\partial p} \right)^\# \tilde{r}.
\]

Note that \( p \) enters the second term only through \( \delta \). Total differentiation with respect to the \( k^{th} \) dimension of the characteristics gives

\[
\left[ \frac{\alpha \partial \tilde{m}_1}{\partial \delta^k} + \left( \frac{\partial \tilde{q}}{\partial p} \right)^\# \frac{\partial \tilde{p}^*}{\partial x_k^j} \right. + \left. \left( \beta_k \frac{\partial \tilde{m}_1}{\partial \delta^k} + \frac{\partial \tilde{m}_1}{\partial x_k^j} \right) + \alpha \frac{\partial \tilde{m}_2}{\partial \delta^k} \frac{\partial \tilde{p}^*}{\partial x_k^j} + \left( \beta_k \frac{\partial \tilde{m}_2}{\partial \delta^k} + \frac{\partial \tilde{m}_2}{\partial x_k^j} \right) \right] = 0
\]

Therefore, the gradient of the equilibrium price function is

\[
\frac{\partial \tilde{p}^*}{\partial x_k^j} = - \left[ \frac{\partial \tilde{m}}{\partial \delta^k} + \left( \frac{\partial \tilde{q}}{\partial p} \right)^\# \right]^{-1} \left[ \frac{\partial \tilde{m}}{\partial \delta^k} + \frac{\partial \tilde{m}}{\partial x_k^j} \right]. \tag{21}
\]

This expression has an intuitive explanation. Suppose that the term \( \left( \frac{\partial \tilde{q}}{\partial p} \right)^\# \) did not exist and there were no reader heterogeneity, i.e. characteristics affected the system only through the mean utility level \( \delta \), and \( \frac{\partial \tilde{m}}{\partial \delta^k} = 0 \). Then, the Jacobian \( \frac{\partial \tilde{p}^*}{\partial x_k^j} \) would be \( -\frac{\beta_k}{\alpha} I \), where \( I \) is an identity matrix. This is because when all readers have the same taste, a combination of an increase in characteristic \( x_{kj} \) by \( \Delta \) and an increase in price \( p_j \) by \( \frac{\beta_k}{|\alpha|} \Delta \) has no impact on utility, hence circulation. Therefore, \( \frac{\partial \tilde{p}^*}{\partial x_k^j} \) would be \( \frac{\beta_k}{|\alpha|} \) if the pricing strategy of \( j \)'s publisher were to keep the circulation at a certain level. And for \( j \)'s opponents, nothing has changed as this combined change of \( x_{kj} \) and \( p_j \) leaves a reader's utility from newspaper \( j \) unchanged. However, a rational price setter can do better. Her objective is to maximize profit instead of keeping up with a circulation level. After an increase in \( x_{kj} \) by \( \Delta \), the publisher can raise the price by more than \( \frac{\beta_k}{|\alpha|} \Delta \). It can keep on increasing the price until the downward-sloping newspaper demand curve determines that any marginal increase in the price will decrease the profit. So, the newspaper publisher does take into account the slope of the demand function \( \frac{\partial \tilde{q}_j}{\partial p_j} \). When the newspaper publisher has multiple newspapers, it also considers the cross-effect among its newspapers, hence, the term \( \left( \frac{\partial \tilde{q}}{\partial p} \right)^\# \) in the expression.

The Jacobian of the equilibrium advertising rate function is

\[
\frac{\partial \hat{r}_{jk}}{\partial x_{kj}} = \frac{\gamma_3}{1 + 1/\lambda_2} \left( \sum_{j' \in J} \frac{\partial \hat{q}_{h}}{\partial p_{j'}} \frac{\partial \hat{p}_{j'}^*}{\partial x_{kj}} + \frac{\partial \hat{q}_{h}}{\partial x_{kj}} \right). \tag{22}
\]
C Data Sources and Definition of Variables

Demand Data on county circulation for newspapers that are members of the Audit Bureau of Circulation (ABC) is from the *County Circulation Report* of ABC. 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 few newspapers have multiple subscription prices. The local price is used. Display advertising rate is the open inch rate measured in dollars per column inch.\(^48\)

Characteristics A newspaper is described by the following characteristics: (1) news hole, the space of a newspaper that is devoted to news, (2) the strength of the opinion-oriented section of a newspaper, (3) the number of staff-bylined stories in a newspaper, (4) frequency of publication, and (5) edition (morning or evening newspaper).

News hole is the difference between total pages and display advertisements. 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 \((\frac{6}{7}, \frac{1}{7})\).

As for the second and third newspaper characteristics, I use the number of staff on opinion sections, such as columnists and editorial editors, and the number of reporters as proxies. Data on these variables 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 this person 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 are similarly defined.

Data on frequency of publication and edition (morning or evening newspaper) is from *E&P*.

Another factor that influences a reader’s utility from a newspaper is the distance of her county to the newspaper’s head county. Information on the head county of a newspaper is gathered from *E&P*. The distance of two counties is computed based on the data of latitude and longitude of county centers provided by the Census Bureau.

The data source and the description of the variables are summarized in Table 11.

\(^{48}\)Therefore, price discrimination in both subscription prices and advertising rates is ignored, albeit for different reasons. I ignore the price discrimination in newspaper prices because most newspapers offer only one price. There is not enough data variation to identify the difference in demand or marginal cost across geographic areas. I ignore price discrimination in advertising rates because of data limitations.

\(^{49}\)Recall that news hole \(x_{1jt}\) is not observable. \(x_{1jt} = n_j f_j - a_j\).

\(^{50}\)Recall that \(D_{1jt} = 1\).
Table 11: Data Description and Source

<table>
<thead>
<tr>
<th>Variable Data Description</th>
<th>Data Description</th>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Newspaper Demand</td>
<td>$q_{jct}$</td>
<td>County circulation</td>
</tr>
<tr>
<td>Display advertising Demand</td>
<td>$a_{jt}$</td>
<td>Annual Display Advertising linage (column inch)</td>
</tr>
<tr>
<td>Price of Newspaper</td>
<td>$p_{jt}$</td>
<td>Annual Subscription Price (1997 $)</td>
</tr>
<tr>
<td>Price of Display Advertising</td>
<td>$r_{jt}$</td>
<td>Advertising Rate (1997 $/column inch)</td>
</tr>
<tr>
<td>Newspaper Characteristics</td>
<td>$x_{2jt}^{49}$</td>
<td>Weighted sum of reporters and correspondents</td>
</tr>
<tr>
<td></td>
<td>$x_{3jt}$</td>
<td>Weighted sum of columnists and editorial editors</td>
</tr>
<tr>
<td></td>
<td>$f_{jt}$</td>
<td>Frequency of publication (issues/52 week)</td>
</tr>
<tr>
<td></td>
<td>$y_{2jt}$</td>
<td>Edition (morning or evening)</td>
</tr>
<tr>
<td></td>
<td>$n_{jt}$</td>
<td>Average pages per issue</td>
</tr>
<tr>
<td>County Distance</td>
<td>$y_{3jct}$</td>
<td>Distance between county $c$ and the head county of newspaper $j$ (100km)</td>
</tr>
<tr>
<td>Owner</td>
<td>Publisher</td>
<td>Bacon</td>
</tr>
<tr>
<td>County Demographics</td>
<td>$D_{2c}^{50}$</td>
<td>% of population over 25 with bachelor’s degree or higher</td>
</tr>
<tr>
<td></td>
<td>$D_{3c}$</td>
<td>Median income (1997 $)</td>
</tr>
<tr>
<td></td>
<td>$D_{4c}$</td>
<td>Median age</td>
</tr>
<tr>
<td></td>
<td>$D_{5c}$</td>
<td>% of urban population</td>
</tr>
<tr>
<td></td>
<td>$D_{6c}$</td>
<td>Number of households</td>
</tr>
</tbody>
</table>

ABC: County Circulation Report by Audit Bureau of Circulation
Bacon: Bacon’s Newspaper Directory
E&P: Editor and Publisher International Year Book
SRDS: SRDS Circulation
TNS: TNS Media Intelligence
D Explanation of the Results in Section 7.2

In this appendix, I provide a detailed explanation for observations (2), (3) and (4) in Section 7.2.

I first show that the advertising profit function is convex in circulation according to the estimates. Advertising profit is the sum of display advertising profit \( (mc_j^{(a)} - r_j) e^{\eta_j} q_j^{\lambda_1} r_j^{\lambda_2} \) and the preprint profit \( \mu_1 q_j + \frac{1}{2} \mu_2 q_j^2 \). Even though the preprint function is concave in circulation \((\hat{\mu}_2 < 0)\), its second order derivative is positive for all newspapers in the sample at the estimates. Note that the estimated elasticity of display advertising demand with respect to circulation is larger than 1 \((\hat{\lambda}_1 > 1)\). Since the circulation profit \( (p_j - mc_j^{(a)}) q_j \) is linear in \( q_j \), the overall profit function is also convex in circulation.

This convexity has two implications. First, because the marginal advertising value of circulation is larger for larger newspapers, a multi-newspaper publisher has an incentive to shift the circulation from its smaller newspapers to larger newspapers. Second, a newspaper has an incentive to increase its circulation when a decrease in quality or an increase in the price of its competitors leads to an increase in its circulation.

Observation (4) is a direct result of the second implication. As Star and Pioneer reduce their overall quality, the Stillwater Gazette, a local newspaper competing with them in Washington County, has an incentive to increase its circulation. Therefore, it decreases its price and improves its quality to achieve that. For example, its incentive for decreasing its price can be seen from the first order condition with respect to price: \( q_j + \left(p_j + \frac{\partial \pi_j^{(a)}}{\partial q_j} \right) \frac{\partial q_j}{\partial p_j} = 0 \), where \( \pi_j^{(a)} \) is the overall advertising profit function. When \( \frac{\partial \pi_j^{(a)}}{\partial q_j} \) increases, it is optimal to decrease price \( p_j \).

I now explain observations (2) and (3) on the two merged newspapers Star and Pioneer. Recall that Star is the larger party. Two forces affect McClatchy’s decision on the quality and prices of these two newspapers. The first force is a quality cross-effect. For example, when Star increases its number of reporters, the circulation of Pioneer falls and hence the profit of Pioneer. The other force is the concern of leaving space in the old quality region for competitors to shift their newspapers there and compete for readers in order to attract advertisers as well.

To understand the quality cross-effect, I plot the profit functions of Star and Pioneer as well as the sum of these two profit functions around the original equilibrium (i.e. pre-merger equilibrium) in Figure 8 and Figure 9. For example, in the left graph of Figure 8, the x-axis represents \( \log(1+\text{news hole}) \) of Star, and the y-axis is profit in million and \((\pi_1, \pi_2)\) represents the profit from Star and Pioneer, respectively.\(^{51}\) Profit is plotted as a function of Star’s news hole when the other dimensions of Star’s quality measures as well as the quality of other newspapers are fixed at the pre-merger equilibrium. In contrast, prices are allowed to fully adjust to a second-stage equilibrium. In the middle and the right graph, profits are plotted as the number of opinion section staff and reporters of Star vary. Figure 9 shows how the profits of Star \((\pi_1)\) and Pioneer \((\pi_2)\) change as the characteristics of Pioneer (in contrast to Star in Figure 8) vary.

The quality cross-effects are shown in the \( \pi_2 \) curve in Figure 8 and the \( \pi_1 \) curve in Figure 9: how the profit from Pioneer varies as the quality characteristics of Star change, and vice versa. The cross-effect curves are downward-sloping in the middle and right graphs of both figures, implying that as one newspaper increases its quality in terms of the number of staff for the opinion-oriented section or the number of reporters, the other newspaper’s profit falls. However, the left graphs in both figures show that the sign of the cross-effect of news hole is not determinate. As mentioned in the intuition for observation (2) in Section 7.2, news hole also affects the marginal cost of increasing

\(^{51}\)For presentational convenience and because only the shapes of the profit curves are relevant for the arguments, I adjust the location of these curves.
circulation. Increasing news hole therefore leads to a higher marginal cost \( mc(q) \) and hence does not always affect the other newspaper’s profit adversely when prices are fully adjusted to an equilibrium. In particular, at the original equilibrium, the cross-effect with respect to news hole is positive. This is consistent with the observation that news hole of both newspapers increases after the merger while the other dimensions of newspaper characteristics decline.

The asymmetric incentive to adjust quality for Star and Pioneer can be seen from comparing the original equilibrium quality characteristics to the argumentum of the total profit function \( \pi_1 + \pi_2 \). After the merger, McClatchy internalizes the cross-effect of Star and Pioneer and maximizes the total profit, the maximum of which is close to the original equilibrium in Figure 8 but noticeably different in Figure 9. For example, the argumentum of \( \pi_1 + \pi_2 \) in the middle graph of Figure 9 is left to the original equilibrium point, indicating that McClatchy can increase its profit by decreasing the number of staff for the opinion-oriented section in Pioneer when other characteristics of Pioneer and the quality of all other newspapers are fixed. In contrast, McClatchy does not have such a strong incentive to adjust the quality of Star, the larger party, as indicated in Figure 8. This difference is due to the convexity of the advertising profit function in circulation. A marginal improvement in the quality of Pioneer leads to a reduction in Star’s circulation, which decreases the profit of Star by a larger margin than the marginal effect of an increase in Star’s quality on Pioneer’s profit. In other words, the cross-effect of Star’s quality on Pioneer’s profit is smaller than vice versa.\(^{52}\)

I have shown that the adjustment under the quality cross-effect only is consistent with the full adjustment at the new equilibrium. This holds for both the direction of adjustment and the asymmetry in adjustment, which means that the cross-effect essentially determines the full equilibrium outcome. This explains observation (2) and (3).

\(^{52}\)The same argument applies to the asymmetry of the price cross-effect, which explains the difference in the price adjustments in Table 6