Product Differentiation: An Investigation of the 1924 U.S. Newspaper Industry

by Yang Xu

An essay submitted to the Department of Economics
in partial fulfillment of the requirements for the
degree of Masters of Arts

Queen's University

Kingston, ON, Canada

August 2015

Copyright © Yang Xu 2015

Acknowledgement

I wish to thank Professor Jaworski for his feedbacks and guidance. Without his help, this project would be a much bigger struggle. I also wish to thank my classmates for their support throughout the program.

TABLE OF CONTENTS

1. Introduction

Government agencies are often concerned with the social welfare of producers and consumers. Consumers derive utility from consumption and producers earn profits when products are sold. The consumer welfare is determined by the quantities sold, demand curve, and market price, and the producer welfare is found by the quantities sold, the supply curve, and the market price (see figure 1). Social welfare is the sum of these two areas. When producers possess market power, they can sell products at a price above their own marginal cost (point A). In these situations, there are consumers who would pay more than the marginal production cost but less than the going market price. The producer refuses to sell because they can earn higher profits by not lowering the price for all existing consumers just to sell one extra unit, and thus some consumers (between A and B) will not benefit from consuming the product.

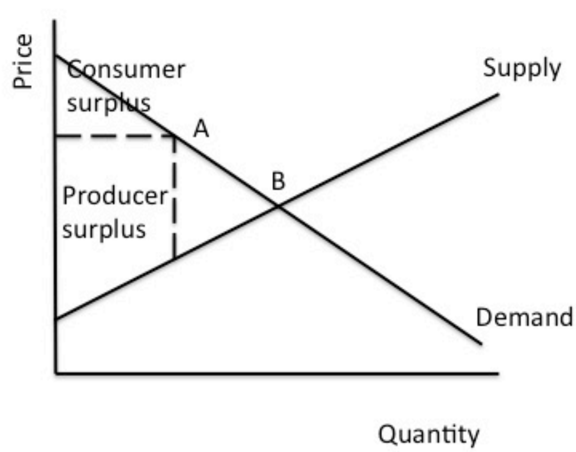


Figure 1: Illustration of market welfare structure.

To sustain high social welfare, one would like the actual trading quantity (point A) to be close to where the marginal demand equals to marginal cost (point

B). Merger guidelines are created to prevent merging firms from raising prices above or lowering quality below premerger levels without incurring a loss of profit.

In certain situations, there are socially benefitting externalities generated by markets that may not be held at optimal levels when firms maximize their profit. Contrasting newspapers, for example, help to provide balanced reporting on political events and their diverse reporting angles cater to consumer preferences. This property has shaped policies pertaining to the newspaper industry. Postal Act of 1792 provided subsidies and the Sherman Act made exceptions for newspaper industry to have joint operating agreements where they could collude on prices for readers and advertisers so long as their contents remain unique (Busterna and Picard 1993). The exception allowed failing newspapers to become financially stable. In many cases, the antitrust agency punishes colluding parties, but an exception was made here to salvage diversity in the media market. These policies illustrate the importance of positive externalities created by newspaper diversity. In this study, I examine the political substitutability perceived by the readers. Using a discrete-choice demand model on a data set containing over 800 U.S. newspapers operating in 300 towns, I investigate how the 1924 consumers differentiate newspapers on the time and political dimension. These data enables me to analyze a period where newspaper was the major source of political information because the radio has yet to gain popularity and the television was yet to be invented.

The existing literature has focused on two main areas. Much of previous merger literature have examined the price effect of proposed mergers, in particular the accuracy of predicted price in merger simulations under different synergy assumptions (Mazur, 2011) and different marginal costs and conduct assumptions (Peters, 2006). As for studies of the newspaper industry, a series of studies by Shapiro and Gentzkow have focused on the choice of political affiliation for entering papers (Gentzkow, Shapiro, and Sinkinson, 2014), incumbent parties' potential influence on the press (Gentzkow, Shapiro, Petek, and Sinkinson, 2015) and effects of firms' entry and exit on voter turnout (Gentzkow, Shapiro, and Sinkinson, 2014). Chiang (2010) used early 2000 data to show that newspaper competition encourages ideology differentiation and draws readers from extreme political views. Mondak (1995) used data on Pittsburgh's newspaper strike in 1992 to study residents' knowledge on local election candidates. See Graber (2000) for a review of the effect media has on politics.

The discrete-choice approach used here is an important technique underlying much of analysis for merger price effect and market demand. Traditionally, to assess whether a proposed merger will have significant price changes, merger simulations are required. In the Canadian Merger Enforcement Guidelines and Competition Act, a routine screen on market concentration is conducted before the cumbersome merger simulation. To calculate market concentration (Herfindahl-Hirschman Index, HHI) for a given market, one needs firms' share of total market sales, profit, or production capacity.

$$H = \sum_{i=1}^{N} s_i^2 \tag{1}$$

The HHI is the summation of the squared shares of all market firms. For example, 30 percent market shares can either be represented as 0.30 or 30 in different formulas. The United States Department of Justice deems markets with an

HHI between 1500 and 2500 moderately concentrated and above 2500 to be highly concentrated. A 200-point increase in HHI for highly concentrated markets is perceived likely to enhance market power. The Canadian Merger Enforcement Guidelines do not have delineated thresholds using HHI.

To calculate market shares, one must first identify the relevant market for the product in question. From the Competition Act, this total market is “the smallest group of products, including at least one product of the merging parties, and the smallest geographic area, in which a sole profit-maximizing seller (a “hypothetical monopolist”) would impose and sustain a small but significant and non-transitory increase in price (“SSNIP”) above levels that would likely exist in the absence of the merger."

These HHI calculation based on premerger data can be misleading as consumers may exit the market leading to different post-merger trading quantities. Furthermore, differences in how the market is defined can influence the values of HHI and affect whether a merger will be allowed. Whole Foods and Federal Trades Commission argued over whether Whole Food was a premium natural and organic supermarket or merely another supermarket. The former market is highly concentrated and the merger will require court hearing, while the latter market definition would not lead to any anti-trust issues. In differentiated product market, this definition could be rather difficult to delineate. Merger simulations on top of HHI would be needed to substantiate court cases.

The analysis conducted in this essay is a simple step that can be added to common merger analysis. Simply looking at HHI will ignore the quality aspect affected by the merger. In this paper, I apply the discrete-choice model to a newspaper data set and use it to consider non-price market characteristics, namely political affiliation. I hope to stress that price coefficient is not the only important criteria in examining potential mergers. Other characteristics may be effectively analyzed in a discrete-choice setup to help define the relevant market.

2. Newspaper Background

The newspaper industry has been providing information to the public since the 1800s, and it precedes the radio and the television as a media form for reporting important events. News can be categorized as "hard news" and "soft news", contrasting political events and entertainment updates. The type of content provided by the media is often catered to their consumers. This is evident in the television market where the less regulated American stations have fewer informational news than the more regulated British stations (Zaller, 1999). Apart from the informative versus entertainment distinction, news can feature local, national, and international contents. Gentzkow and Shapiro (2008) mention that more competitive firms have less informative content on local events than monopolistic ones. When conducting market analysis on the media industry, the researcher must take these product characteristics into account while analyzing the price effects.

Reporting is often subjective and diverse sources will give readers a more complete perspective. Newspapers have their own bias when reporting events, for example New York Times and Fox News may emphasize different aspects of the same story or cover different stories in service of a particular agenda (Alterman, 2003; Coulter 2003, Goldberg, 2003). When reporting scandals, the political ideology of the paper will often determine the connotation used in the headlines. While one source may use strong words such as fraud, bribery, and rob, another paper may appeal to a different set of evidence to claim that the politician is vilified and the stories are merely accusations. By having a diversified reporting angle, consumers may benefit from hearing both sides of the event and make their own judgements. Furthermore, the lack of competition can lead to information being suppressed. Without contrasting ideology, a Republican newspaper may be reluctant to break the story on a Republican scandal and leave out important details. Consequently the public would remain less well informed. In the Credit Mobilier case, the presence of Democratic sources led to more detailed reporting by Republican firms (Gentzkow, Glaeser, and Goldin, 2006). Thus having ideologically diverse set of newspapers will keep the readers better informed even if they only subscribe to a single source.

In terms of anti-trust regulation, the traditional market definition needs to be carefully examined. According to Gentzkow and Shapiro (2008), there are two markets in question for media outlets. The traditional subscriber's market is one where firms in close geographical proximity compete for readers. The second market is an information market that exists because news can be disseminated quickly. A news blog would be considered part of the same information market as the New York Times and the Boston Globe, even though they are geographically separate. It is an online news source that first broke the story between Monica

Lewinsky and Bill Clinton (Stewart, 1999), but the story later became wide spread across all media outlets. The market definition Herfindahl index would not be very informative when assessing the information market (Gentzkow and Shapiro, 2008).

Many newspaper firms in the early 1920s had declared their political affiliation and in return governments offered patronage in the form of jobs, contracts, or subsidies (Gentzkow, Shapiro, and Sinkinson, 2014). Those that never declared affiliations often focused on non-political contents. 75 percent of the papers that focus on commercial, financial, legal, or trade matters never declared a political affiliation (Gentzkow, Shapiro, and SInkinson, 2014). In certain situations, political parties even paid money to introduce new papers, bailing out existing ones, and withdrew support if papers did not follow the party ideology. Without surprise, about 20-40 percent of the newspaper coverage focused on political news (Baldasty, 1992), and previous studies found that consumption of newspapers had a positive effect on voter turn out (Stromberg, 2004; Gentzkow 2006; Snyder and Stromberg, 2008). In the U.S., despite lucrative government printing contracts (which can account up to 50 percent of some newspaper firms' revenue) and patronage positions, there is little evidence that the incumbent government had strong influences on the newspaper market when one examines circulation shares and entry data (Gentzkow, Shapiro, Petek, and Sinkinson, 2015).

Before the introduction of radio and television, newspaper was the main source of political information, and today newspaper continues to be an important source of local news. According to Mondak (1995), Pittsburgh newspaper strike significantly affected the public's knowledge regarding the congress candidates while it had no effects on presidential candidates. This partly speaks to the dissemination property of the news industry, where the information reported by other cities' news outlets can easily spill over to Pittsburgh.

On the readership side, 15 percent of reader households purchase multiple products and have a strong preference for papers sharing the same political affiliation. Consumers tend to read papers that agree with their prior beliefs, an idea that is substantiated by both psychology (Lord, Ross, and Lepper, 1979; Nisbett and Ross, 1980) and econometric data (Gentzkow and Shapiro 2006, 2007). Mullainathan and Shleifer (2005) and Gentzkow and Shapiro (2006) even suggested models where media distort information to cater to their consumers' preferences.

As another source of potential differentiation, some newspapers are morning editions while others are evenings. In markets with more than one newspaper, the entrant usually does not occupy the same time and political dimension as the incumbent, and it is certainly difficult to take the dominant position away from the existing paper (Hamilton, 2006). When an entrant comes into the market, 86 percent of the readers come from people who are either ordering an additional paper or did not subscribe previously.

Perceived reputation also differentiates newspapers. Certain firms will commit more effort to report news in a timely and accurate manner, and their coverage will be more comprehensive. In monopolistic markets, papers will either exert little effort to keep a well regarded reputation due to lack of competition or if they do commit effort, they would recover greater rents (Gentzkow and Shapiro, 2008). Interestingly firms insulated from competitive pressure can often be more informative, as in the case of the BBC (Prat and Stromberg, 2005), while firms in competitive environments can be limited in terms of reporting and editorial assets (Zaller, 1999). Newspapers had significant market power in the 1920s. For an average copy, \$4.69 of the revenue came from readers and \$14.19 from advertisers, while the marginal cost was only \$10.09. During that time more competitive markets had higher prices (Gentzkow, Shapiro, and Sinkinson, 2014).

3. Discrete-choice Model of Demand

The demand estimation procedure follows the discrete-choice model for differentiated product market proposed by Berry (1994) under the assumption of perfectly competitive and static equilibrium. His procedure incorporates a utility function that depends on consumer specific preferences and the characteristics of purchased goods. Given that products have different traits and prices, consumers choose the good that gives them the highest utility.

In Berry's approach, market demand is derived from the utility of individual consumers, which is represented by the following equation:

$$u_{ijt} = x_{jt}\tilde{\beta}_r - \alpha_r p_{jt} + \xi_{jt} + \varepsilon_{ijt} \tag{2}$$

Consumer i's utility for buying the product j in market t is a function of observable characteristics (x_{jt}), price (p_{jt}), characteristics unobserved by the econometrician (ξ_j), and an independent and identically distributed (iid) error term (ε_{ijt}) representing consumer specific variations around the average utility amongst all consumers for good j (δ_j). Parameters to be estimated are β_r and α_r. The taste parameter β can be consumer specific, and we can modify the model to accommodate. For characteristic k:

$$\beta_{ik} = \beta_k + \sigma_k\zeta_{ik} \tag{3}$$

$$\nu_{ij} = \sum_k x_{jk}\sigma_k\zeta_{ik} + \varepsilon_{ij} \tag{4}$$

$$u_{ij} = x_j\beta - \alpha p_j + \xi_j + \nu_{ij} \tag{5}$$

If individual consumer heterogeneity ν_{ij} is only included in the model through ε_{ij} and ε_{ij} are iid, then the model collapses to the simple logit model.

Having the information on observables, prices, and utility maximizing choices, one can formulate a set of unobservables that allows consumer to purchase product j.

$$A_j(\delta) = \{\nu_i \backslash \delta_j + \nu_{ij} > \delta_k + \nu_{ik}, \forall k \neq j\} \tag{6}$$

A market share function predicting the share of consumers that purchase each product can be constructed by integrating the unobservables over the set $A_j(\delta)$, given the probability density function $f(\nu, x, \sigma)$. f_j denotes the market share function of product j.

$$f_j(\delta(x,p,\xi),x,\theta) = \int_{A_j(\delta)} f(\nu,x,\sigma_\nu)d\nu \tag{7}$$

Berry (1994) shows that there is a one-to-one relationship between the actual market share and the mean utility: $s_j=f_j(\delta)$. So one could map the observed shares uniquely to the mean utility for each product. One only needs the share information to calculate the mean utility values—the dependent variable of the linear regression procedure used to find price and characteristic coefficients. A set of sufficient conditions exists for the one-to-one relationship: the share function is differentiable with respect to the true mean utility everywhere; the own derivatives are strictly positive and cross derivatives are strictly negative; for all characteristics, the population density function for consumer characteristics, f(ν,x), needs to be strictly positive and continuous for all consumer characteristics (or called the consumer taste parameter in equation (4)). A detailed proof of this relationship can be found in the appendix of Berry's (1994) paper.

3.1 Logit Model

In the simple logit model, the share formula is given by:

$$f_j(\delta) = \frac{e^{\delta_j}}{\sum_{k=0}^{N} e^{\delta_j}} \tag{8}$$

$$\text{where } \delta_j = x_j\beta - \alpha p_j + \xi_j \tag{9}$$

There is often correlation between the price and the unobserved term but in the share function they are related non-linearly. Unobservables are included because consumers value newspaper quality and content variety, which are not observed by the econometrician. The traditional instrumental variable approach cannot be directly applied. Assume that unobservables are known, one can transform the market share function to depend only on δ_j, the average consumer utility for good j.

$$\delta_j = \ln(s_j) - \ln(s_0) = x_j\beta - \alpha p_j + \xi_j \tag{10}$$

$\delta_0 \equiv 0$ as normalization. The consumer's utility from purchasing the outside good only depends on their specific taste parameter, $u_{io}=v_i$. This outside good acts as an alternative choice to purchasing any of the newspapers amongst the differentiated products. The outside good is needed to account for uniform price increase in the industry. Without it, consumers will be forced to choose amongst the more expensive products.

Though analytically simple, the substitution pattern implied by the logit model is rather restrictive. As in Ackerberg and Crawford (2005), the own and cross product elasticity can be calculated as below. Appendix section of this essay shows the derivation for the more complicated nested logit elasticities.

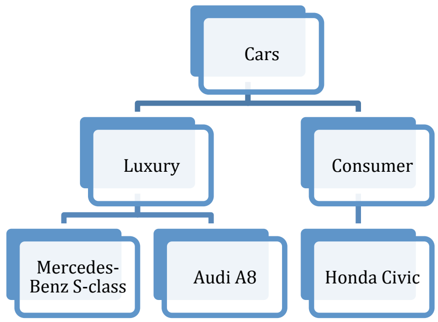
$$\frac{\partial s_{jt}}{\partial p_{jt}} = -\alpha s_{jt}(1 - s_{jt}) \tag{11}$$

$$\frac{\partial s_{jt}}{\partial p_{kt}} = \alpha s_{jt} s_{kt} \tag{12}$$

Only the market shares matter when calculating cross product elasticity, but this creates restrictions in substitution patterns. For example, a Mercedes-Benz S-class, Audi A8, and Honda may have the same market share in car sales, but people would view BMW and Mercedes as luxury vehicles and consider them as closer substitutes. The nested logit model allows for these kinds of differences in substitution pattern.

3.2 Nested Logit Model

In Berry (1994), the nested logit model was discussed to allow for correlation in consumer taste for products within the same nest. The multinomial logit model assumes that the choices are independent of irrelevant alternatives, meaning we cannot have groups of products more similar to each other than rest of the choices. If two goods belong to the same nest, then the ratio for the probability of selecting good 1 over the probability of selecting good 2 ($Prob_1/Prob_2$) is the same for both the logit and the nested logit model. If two goods belong to different nests, then the ratio is different.



In the example used in section 3.1, I had a nest structure for consumer goods. The relevant price and characteristic coefficients can be estimated using either maximum likelihood or sequential logit estimation. In the first step of the sequential technique, I estimate the probability of choosing the Audi when consumer is deciding between luxury vehicles. Then I estimate the probability of choosing luxury vehicles over consumer vehicles. The W represents the set of characteristics that is unique to each car in the nested group and not shared between them. Y_i are the characteristics not shared between nests but shared within a nest (Greene, 2011).

$$P(Audi|Luxury) = \frac{e^{W_{Audi}}}{e^{W_{Audi}+e^{W_{Mercedes}}}} \tag{13}$$

$$P(Luxury) = \frac{e^{W_L+\sigma IV_L}}{e^{W_L+\sigma IV_L+e^{W_C+\sigma IV_C}}} \tag{14}$$

$$IV_L = ln \sum_{i \in L} e^{Y_i} \tag{15}$$

These nested logit models are widely used in empirical literature. Analysis of the airline industry often include all air travel itineraries as the nested good and other travel methods as the outside good (Berry and Jia, 2006; Craig Peters, 2006).

In this study, I use morning and evening editions as nests to examine consumer substitution preferences. Utility function for nested logit is given as:

$$u_{ij} = x_j\beta - \alpha p_j + \xi_j + \nu_i(\sigma) + (1-\sigma)\varepsilon_{ij} \quad (16)$$

ν_{it} differentiates the different nests. It is a random taste parameter. σ varies between 0 and 1, where σ=0 turns the formula into its multinomial logit form.

The share of consumers purchasing good j is represented by equation

$$f_j(\delta,\sigma) = f_{j|g}(\delta,\sigma)f_g(\delta,\sigma) = \frac{e^{\delta_j/_{1-\sigma}}}{D_g^{\sigma}[\sum_g D_g^{1-\sigma}]} \quad (17)$$

$$D_g = \sum_{j\in g} e^{\delta_j/_{1-\sigma}} \quad (18)$$

$f_{j|g}$ is the nested share for good j. After inverting the share function, I get a simple formula that I can use to estimate the price and characteristic coefficients. I simply regress δ_j on x_j, p_j, and $\ln(s_{j|g})$.

$$\delta_j = x_j\beta - \alpha p_j + \xi_j + \sigma\ln(s_{j|g}) \quad (19)$$

The nested models allow for different substitution patterns for goods that are inside a subgroup. The own-price elasticity and cross-price elasticity can be calculated for the nested logit models using equations (20), (21), and (22) below. The characteristic elasticity formulas are the same except β replaces α. (see Appendix for derivations)

$$\frac{\partial s_{jt}}{\partial p_{jt}} = -\alpha s_{jt}\left(\frac{1}{1-\sigma} - \frac{\sigma}{1-\sigma}s_{jt|g} - s_{jt}\right) \quad (20)$$

$$\frac{\partial s_{jt}}{\partial p_{kt}} = \alpha s_{kt}\left(\frac{\sigma}{1-\sigma}s_{jt|g} + s_{jt}\right) \quad (21)$$

$$\frac{\partial s_{jt}}{\partial p_{kt}} = \alpha s_{jt}s_{kt} \quad (22)$$

The nested logit model has two ways of calculating the cross-product elasticity. If the goods belong to the same nest, equation 4 will be used, otherwise the elasticity is the same as the logit elasticity. As one can see, the nested parameter σ is the difference in the elasticity formula. If $\sigma=0$, the elasticity collapses into the regular logit formula. The substitution pattern in nested logit model is based on preselected grouping of products, which is decided by the econometrician before estimation. The random coefficient model has substitution pattern depending on a random coefficient parameter that is continuous. Nonetheless, nested logit elasticity is more flexible than the multinomial logit but is still restricted compared to the full random coefficient model. However, the more flexible model will be much more computationally intensive, especially when calculating the predicted market shares, as it requires integration. The full random coefficient model is beyond the scope of this essay.

3.3 Instrumental Variables

Price can be endogenous with unobserved price characteristics such as product quality, content variety etc. Because consumers may value content variety in their product, and they are willing to pay a higher price for it. In linear regression model, $y = x'\beta + u$, if the regressor is correlated with the error term, then the regression coefficient β will be inconsistent. β will not only reflect the impact of regressor x on the variations in y but also some of the variations in u that is correlated with variations in y. The latter effect is caused by the correlation between the regressor and the error term. To consistently estimate the coefficients, I need instruments that are correlated with the regressors but uncorrelated with the error term. With these instruments, I can generate predicted values of regressors and use these to construct the consistent β coefficients using the two-stage least squares method. In the share function, the unobservable characteristics and price term are non-linear in specification, and I cannot use instrumental variable for non-linear terms. Berry's (1994) method transform the non-linear share function into a linear form which I could apply instrumental variable estimation.

In my study, I will need instruments for prices. Input prices, price and x_k from other products are appropriate choices (Berry, 1994). In situations where cost data are unavailable, variables that affect mark-up and marginal cost can both serve as instruments (Berry, Levinson, and Pakes, 1995). To exploit how differentiated product firms take other firms characteristics into account to set their own prices, Fan (2012) used demographic information from competing firms' markets as instruments for price. In my available dataset, I do not have extensive amount of instruments because the original paper was intended to focus on the study of political factors that affect entry and affiliation decisions of newspapers.

To simplify the model, I assume that one individual chooses only one paper per time period. Continuous or multiple purchases can be found in Dubin and McFadden (1984) and Hendel (1999). Fan (2012) examines multiple newspaper purchases by a single household and takes diminishing marginal utility into account.

3.4 Supply side

In Berry's (1994) paper, there are assumptions made regarding the supply side of the model. Firms are price setters and the equilibrium in the market is Nash in price. Firms wish to maximize their profits.

$$\pi_j\ (p,z,\xi,\omega_j,\theta) = p_j M s_j\ (x,\xi,p,\theta_d) - C_j\ (q_j,w_j,\omega_j,\gamma) \tag{23}$$

γ is an unknown vector. M is the size of total market. I can also split the cost term into fixed and marginal costs. For a multiproduct firm

$$\pi_F = \Sigma_{j\in F}\ (p_j - mc_j) M s_j\ (p) - C_j \tag{24}$$

The profit maximizing condition in algebraic and matrix notation are

$$s_j(p) + \Sigma_{r\in F}\ (p_r - mc_r)\ \partial s_r(p)/\ \partial p_j = 0 \tag{25}$$

$$s(p) - \Omega(p)(p-mc) = 0 \tag{26}$$

When marginal cost is needed to calculate the post merger equilibrium, many empirical papers estimate the marginal cost using this first order condition.

$$mc = p - \Omega\ (p)^{-1}\ s(p) \tag{27}$$

Ω matrix will be calculated using the specification of the share function. Post merger price can be calculated using marginal costs recovered from premerger data. Since I do not have large amounts of firm characteristics to conduct a proper supply side analysis, I shall focus on the demand side for this paper.

4. Data

The data set was initially used by Gentzkow, Shapiro, and Sinkinson (2014) to examine ideological diversity in the U.S. newspaper markets. Whereas they examined political affiliations and newspaper entry, I focus on the discrete-choice demand estimation and substitution patterns.

The original data from the paper come from a variety of sources, including newspaper directories, which reports affiliation and circulation prices, and auditing agency which had data on town level circulations. Potential noise exists as these values are self-reported by the newspaper firms. The original authors gathered cost and revenue data for a subset of papers from income statements provided by the Inland Daily Press Association. There is also a panel data set that has newspaper name, political affiliation, city, and subscription price of daily newspapers from presidential years between 1869 and 1924. Only the 1924 data were used for analysis as I only had instruments for the 1924 subscription prices. From these available data sets, I gathered newspaper variable cost, per copy advertising revenue, political affiliation, time of day the paper was circulated, subscription price, and quantities.

For the original paper, they modeled entry of newspapers into various towns, and therefore needed town-level circulation data. I used headquarter data instead because there was better proxy available for market size in headquarter cities. Newspaper can circulate in several neighboring towns, so the market share is calculated by the number of subscriptions divided by the total population for the trading area. For tractability, I assume that market share is uniformly distributed across the trading area and I can calculate outside share for each town from the individual papers' market shares. Certain markets have multiple newspapers present and provide political diversity for their readers. Papers can have several subscription prices, giving readers a choice to pay for each day of the week or just certain days. I use the price for the most comprehensive subscription package issued. In my data, the political affiliation is a binary parameter where 0 equal to Democratic affiliation and 1 equal to Republican.

Table 1: Edition and political affiliation of estimation sample

	Democratic	**Republican**	**Total**	**Omitted samples**
Morning	73	77	150	89
Evening	118	257	375	191
All Day	2	5	7	6
Total	193	339	532	286

Note: The omitted samples column includes observations used to calculate market share but excluded from the regression for reasons described in the main text.

Table 2: Average share values for various subgroups.

	Average share (standard error)
Morning	0.1065 (0.05505)
Evening	0.1140 (0.5443)
All Day	0.1433 (0.06109)
Republican	0.1213 (0.05354)
Democratic	0.09634 (0.05348)

Notes: The data used in rows 1-3 is the full sample and Democratic and Republican shares are calculated from papers that have declared their affiliations.

In the full sample, some observations did not have Record ID to match newspaper circulation values with its price and cost variables. Those papers were, however, still included to calculate the appropriate market shares. It is only during the regression analysis that these observations were dropped. Certain newspapers' political affiliations were independent and in Gentzkow, Shapiro, and Sinkinson (2014), they found that the omission of these papers did not affect the conclusions. As mentioned before, the papers that focused on topics dealing with finance and commercial contents often did not declare political affiliations. To gauge the validity for excluding observations, I calculated cross elasticity in both the estimation sample and the full sample, summarized their respective average subscription prices, and ran an OLS regression using subscription price as the sole regressor for comparison. In terms of edition for the newspapers, there are more evening than morning papers, in both the full sample and the regression sample. Few all day newspapers also exist.

Table 3: Newspaper market environment

Total competition	1	2	3	4	5	6	7	8	26	Total
Republicans	172	146	20	0	0	0	0	0	1	339
Democratic	49	137	6	0	0	0	0	0	1	193
Total	221	283	26	0	0	0	0	0	2	532
Excluded Sample	38	45	79	48	25	6	13	8	24	286

Notes: New York City is the only town with 26 total daily newspapers. Total competition represents the total number of newspapers in a given newspaper town, including the newspaper being examined.

I had 532 newspapers that had all the data required for the full regression specification. There are 339 Republican papers and only 193 Democratic. I had to drop samples from the set I used to calculate market shares because the newspaper had either independent or have unknown affiliations. The variable total competition reflects the total number of newspapers exist in a town or city before I dropped observations due to insufficient data. In table 3, I tallied the newspapers by size of market competition and political affiliation. 504 out of 532 newspaper in our regression sample were either monopoly or duopoly papers. For computational reasons, the Gentzkow, Shapiro, and Sinkinson (2014) excluded 52 towns with 10 or more circulating dailies in their analysis. I suspect that is reason for the missing political affiliations.

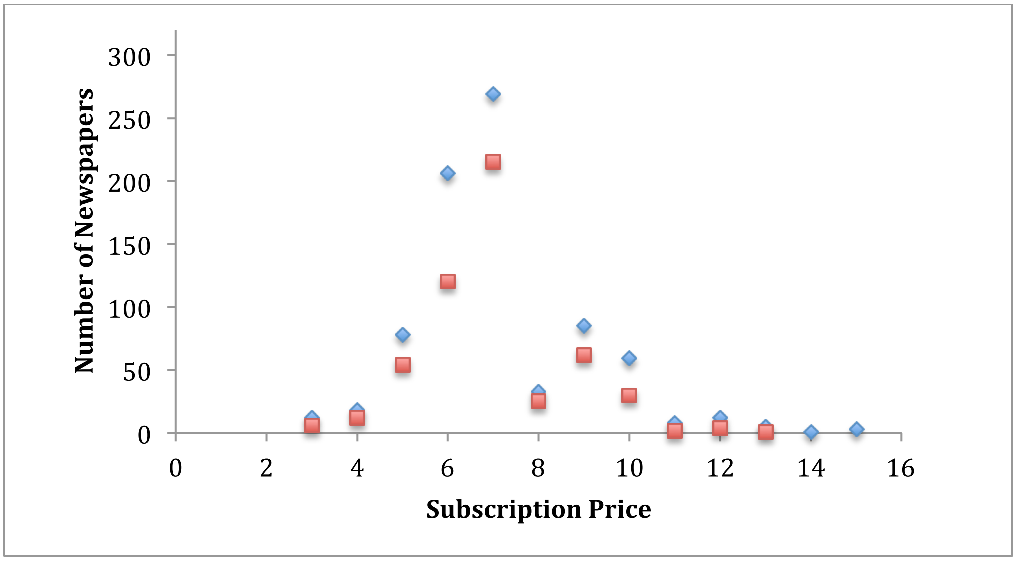


Figure 1: Scatter plot of subscription price from demand estimation sample (square) and full share calculating sample (diamond).

The figure above shows the subscription price distribution for the estimation and the full sample. The only drastic difference is seen at the price of $6, while the rest of the price ranges have similar ratio of papers.

Table 4: Mean price and cost values by edition and affiliation.

	Subscription price	Advertising revenue	Marginal Cost	Fixed
Morning	7.78 (1.64)	15.69 (5.59)	10.19 (3.30)	7.96 (4.01)
Evening	7.12 (1.49)	15.53 (5.09)	9.74 (3.02)	8.53 (4.56)
All day	7.48 (2.41)	18.80 (3.93)	12.02 (3.69)	9.95 (2.57)
Democratic	7.40 (1.63)	15.88 (5.47)	10.24 (3.37)	8.24 (3.67)
Republican	7.27 (1.55)	15.47 (5.08)	9.72 (2.96)	8.47 (4.76)
Unknown affiliation	7.62 (2.34)	No data	No data	No data

Note: Full sample is used to calculate the mean values. Standard errors are included in brackets. Observations with unknown affiliations also lack advertising and cost data.

In Gentzkow, Shapiro, and Sinkinson's (2014) summary statistics, 55% of all newspapers were Republican and the average Republican vote share in circulating towns was 51%. For the sample I used, 63% of the papers are Republican. Initial data had 818 newspapers. In Gentzkow, Shapiro, and Sinkinson (2014), they counted 1338 total newspapers but not all observations had newspaper names or Record ID. Thus I cannot match up all papers with the cost data. I discovered 1679 unique Record IDs, but certain newspapers had morning and evening editions each with a unique ID number, which may explain the discrepancy with the 1338 total count. I make the assumption that the regression coefficient from the 532 papers sample and the full 818 sample are the same, so I can analyze the time dimension elasticity on both the small and the large sample. This assumption may lead to problems in the results section, and I am aware of potential problems and shall be cautious in my conclusions.

5. Results

5.1 Regression analysis

Table 5: Demand estimates for logit model.

Demand variables	**OLS**	**OLS**	**IV**	**IV first stage**	**IV**	**IV first stage**
Political affiliation	______	0.2737 **	______		0.2890**	
		(0.0571)			(0.0626)	
Constant	-2.236**	-2.434	-3.033**		-3.601**	
	(0.1285)	(0.1319)	(0.5968)		(0.6373)	
Price	0.0158	0.0190	0.1253		0.1777*	
	(0.0172)	(0.0195)	(0.0816)		(0.0859)	
Political affiliation				______		-0.0668
						(0.1442)
Advertising revenue				-0.0416		-0.0408
				(0.0246)		(0.0246)
Variable cost				0.1274**		0.1255**
				(0.0424)		0.0425
Political competitors				0.3448*		0.3413*
				(0.1526)		(0.1539)
Constant				6.258**		6.311**
				(0.2894)		(0.3201)
R-squared	0.0016	0.0461	0.0386		0.0391	
Adjusted R-squared	-0.0003	0.0425	0.0331		0.0317	
Observations	531	531	528		528	

Notes: The sample only includes the observations with Republican or Democratic as declared political affiliation. Dependent variable is $\ln(s_j)-\ln(s_o)$. First stage results for 2SLS are also reported. Robust standard errors are included brackets. * indicates 95% level of significance. ** indicates 99% level of significance

Table 6: Demand estimates for time based nested logit model.

Demand variables	**OLS**	**OLS**	**IV**		**IV**	
Political affiliation	______	0.2745 ** (0.0570)	______		0.2909** (0.0628)	
Nested share	0.1981* (0.0879)	0.2016* (0.0887)	0.1222 (0.1188)		0.0945 (0.1312)	
Constant	-0.218** (0.1573)	-2.38** (0.1560)	-3.101** (0.5716)		-3.657** (0.6063)	
Price	0.01167 (0.0204)	0.0148 (0.0199)	0.1366 (0.0772)		0.1868* (0.0807)	
Political affiliation				______		-0.0579 (0.1425)
Nested share				0.7629** (0.2791)		0.7611 (0.2797)
Advertising revenue				-0.0415 (0.0240)		-0.0408 (0.0241)
Variable cost				0.1327** (0.0411)		0.1311** 0.0413
Competitors				0.3693* (0.1544)		0.3662* (0.1557)
Constant				6.2621** (0.2900)		6.3087** (0.3207)
R-squared	0.0095	0.0542	0.0571		0.0574	
Adjusted R-squared	0.0057	0.0488	0.0499		0.0484	
Observations	531	531	528		528	

Notes: Nested logit version of table 5. Time nested group share is included as an additional regressor. All day papers are excluded from analysis.

Table 5 and 6 report the parameters for the logit and the nested logit models with robust standard errors. Robust standard errors account for potential heteroskedasticity in the error term across various prices. The unobserved quality may vary more when the prices are high. Higher prices are reflective of either better quality or higher market power, and different papers will have different reasons for their mark-up. The dependent variable in these regressions is the mean utility level of good j, $\delta_j=\ln(s_j)-\ln(s_o)$. In Gentzkow, Shapiro, and Sinkinson's (2014) paper they ignored price unobservables in order to concentrate on the effects of political variables. Since the positive price coefficient did not affect the qualitative interpretations on the political and entry variables, the authors skipped over the interpretation of the price coefficient. The positive coefficient conflicts with standard theory because it suggests that if the price of a product is higher then it becomes more attractive to readers. For example, in column 2 of table 6, a dollar increase in subscription price would lead to 0.0148 percent increase in subscription shares to outside share difference. For the OLS regressions, all the price coefficients are positive, but none of them are statistically significant. I fail to reject the hypothesis that price has no impact on market shares. For the instrumental variable regressions, the price coefficients are both positive and significant. Gentzkow and Shapiro (2008) discussed the situation where product quality is reflected in prices. A monopolist who puts in more effort in writing stories will charge higher prices for better stories, and the positive coefficient may reflect these unobserved qualities. Furthermore, my estimation sample consists mostly of monopoly markets. For the markets where I had full data (quantity, price, instruments, and political affiliation), three quarters of these markets had either one or two newspapers. Firms cannot steal readers from their rivals by lowering prices like they could in competitive markets. The quality of the contents and degree of catering to consumers' preferences will play a bigger role than price in determining market share. This situation agrees with the endogeneity problem Berry (1994) discussed for differentiated product market. The correlation between price and unobservable characteristics, such as quality or service, will give positive coefficients.

With the available data, I estimated an instrumental variable regression to correct for the endogeneity problem. Since newspapers are a two sided market, readers purchase newspapers and advertisers pay for advertising space. In a few papers, this is explicitly modeled by a supply side regression; here I simply use the advertising revenue as an instrument for price. I expect higher advertising revenue will translate into a lower subscription price. Earnings from advertising and subscription are both used to cover operation costs, and advertising accounts for a greater share of the revenue as discuss in section 2. Lowering subscription prices can recruit more readers to make advertising in the paper more appealing. In the opposite situation, if a firm already has a high advertising revenue, they have the incentive to keep prices low and maintain a large readership base. This intuition is supported by the negative sign in the first stage of two stage least square regression. The variable cost term had a positive coefficient as instrument on price, and this is expected because producers require compensation for higher cost, which is at least partially passed onto consumers. I also tried to model newspaper quality using fixed cost data, but fixed cost data was highly insignificant (as presented in the Appendix). The variation in rent and price index across cities perhaps explains the variations in fixed cost better than editors' real salaries.

The political affiliation variable examined in these regressions was not present in Gentzkow, Shapiro, and Sinkinson's (2014) paper. They focused more on the substitutability of different political affiliations and used a much more elaborate set of regressors for political affiliations. Accounting for multiple newspapers in the consumption bundle, they were able to identify the average utility of consuming same or different affiliation papers as well as the substitution between same and different affiliation papers. I can relate to their findings by examining my calculated elasticity values, which is discussed later on.

In my regression tables, the political affiliation terms are significant at the 1 percent level. From table 5 column 5, given that a paper switches affiliation from Democratic to Republican, the firm's market share relative to the outside share will increase by 0.289 percent. Since there is a higher vote share for Republicans and psychology research supports preferences for likeminded news (Gentzkow and Shapiro 2006, 2007), the positive coefficient reflects the population preference for Republican paper. In the first stage of IV regressions, political affiliation had no impact on price itself.

In the nested logit regressions, the nested coefficient is positive but weakly so. Using the OLS with political affiliation column in table 6, if the share of newspaper in its edition nest group were to increase by 1%, its share relative to outside share will increase by 0.20%. Higher nest group share for a paper is correlated to its higher overall share, and this agrees with intuition because the nested group is merely a subset of the whole group. More dominant position in the subgroup translates into more dominant position in the local market. This nested coefficient also represents the degree of correlation between the unobserved terms within the nest. Here our value is not significantly different from zero, there is no evidence to support that readers view morning and evening newspapers as different nested goods even though news are published in different time slots. To verify this claim, I ran another regression using only observations with two or more newspapers in the nested groups. Results are presented in table 7 and here the nested coefficient is highly significant and has a large value. This result suggests that consumers do perceive news as closer substitutions if they share the same time slot. The contrasting results suggest that sample composition matters for nested logit analysis. If most of the nests consist of monopoly markets then the coefficient would be insignificant.

Table 7: Sample composition analysis for time based nested logit model

Demand variables	**Competitive markets**		**All markets**	
Political affiliation	0.1272		0.2909**	
	(0.1083)		(0.0628)	
Nested share	0.7888**		0.0945	
	(0.2227)		(0.1312)	
Constant	-1.932**		-3.657**	
	(0.7014)		(0.6063)	
Price	0.03152		0.1868*	
	(0.1019)		(0.0807)	
Political affiliation		-0.1833		-0.0579
		(0.3894)		(0.1425)
Nested share		-0.1582		0.7611
		(0.9076)		(0.2797)
Advertising revenue		0.04603		-0.0408
		(0.4498)		(0.0241)
Variable cost		-0.07337		0.1311**
		(0.07797)		0.0413
Competitors		1.0698**		0.3662*
		(0.3952)		(0.1557)
Constant		5.4253**		6.3087**
		(1.0606)		(0.3207)
R-squared	0.1259		0.0574	
Adjusted R-squared	0.0719		0.0484	
Observations	87		528	

Notes: First stage results are presented in columns 2 and 4. Only markets with more than 1 paper belonging to the same time slot is included in column 1 regression. Their nested share is less than 1. Column 3 and 4 are reproduced from table 6 for comparison purposes. Robust standard errors are included in brackets. * indicates 5% significance level. ** indicates 1% significance level.

The constant terms in my regressions are always negative and significant. This implies that given the price is 0 and given Democratic as the political affiliation, the share of good j is less than the outside share. Table 2 shows the average share values and they are close to 0.1. With a few papers in each market, s_j will be less than s_0.

5.2 Elasticity Analysis

I organized various elasticities by political affiliation and editions. For own political elasticity, Republican papers are more elastic to changes in political views than Democratic, though the values are within one standard deviation of another. As in row 1 of table 8, if the Republican paper changes their view to Democratic, they would expect to lose 0.028% of market share where as the Democratic switching view would gain 0.023%. This result agrees with vote share data where the Republicans are favoured and the regression coefficient value for political affiliation. For own political elasticities, the specification of the model, whether OLS, IV, or Nested, did not give drastically different results.

Table 8: Average own political elasticity summaries by political affiliation

	Democratic	**Republican**	**Total**
OLS	0.02305 (0.01088)	0.02839 (0.01073)	
IV	0.02434 (0.01149)	0.02999 (0.01133)	
Nested OLS	0.02357 (0.01099)	0.02903 (0.01077)	
Nested IV	0.02469 (0.01158)	0.03042 (0.01138)	
Observations	193	339	532

Notes: Sample included here is the demand estimation sample. Political affiliation is a categorical variable as described in the data section. Column 1 represents changes in share if newspaper changes from Democratic to Republicans, column 2 for

Republicans to Democratic. Since they go in opposing directions, I excluded the average total sample column. Standard errors are included in brackets. Elasticities are calculated according to formula given in the model section.

For cross-political elasticities, Republicans has higher elasticities than Democratic. Between Republican newspapers, there is more gain when other firms changing their political affiliation to Democratic. Since the data set has many more Republican newspapers, this could be reflective of their higher competition level. It is interesting to note that the time nested models gave higher cross elasticities than non nested counter parts, even though the nest are time based not political affiliation based. This suggests that the variables used to specify the nest structure can be different from the variable of interest. I did not have the full political affiliation data to construct ideology nests, but time-based nests gave me different political elasticities than OLS based nests. When papers are separated into nests, there are fewer ideological competitors in the subgroup. Perhaps the changes in ideology are magnified for the subgroup of readers that are affected. This difference between nested and multinomial logit models is not significant, but the pattern is uniform across political affiliations and OLS versus IV models.

Table 9: Average cross-political elasticity summaries by political affiliation

	Democratic	**Republican**	**Mixed**
OLS	0.002749	0.003230	0.002993
	(0.00360)	(0.00197)	(0.00241)
IV	0.002904	0.003411	0.003161
	(0.00380)	(0.00208)	(0.00255)
Nested OLS	0.002758	0.004245	0.003933
	(0.00361)	(0.00293)	(0.00281)
Nested IV	0.002922	0.003873	0.003589
	(0.00383)	(0.00240)	(0.00263)
Observations	31	45	78

Notes: These elasticities are only available for cities with more than one circulating daily newspaper. Standard errors are included in brackets. The sample here only includes papers that have declared their political affiliations.

In the time separated own elasticity groupings, the morning and evening editions had virtually the same elasticities, though the evening elasticities are slightly higher. From table 2, both editions had similar shares, it is reasonable for them to have similar elasticities. The all day papers had a noticeably higher value than the morning or evening only editions. Since all day papers sell during both morning and night, I expect them to sell more papers in a particular market. In table 2, I do show that they have greater market shares. When the political affiliation changes, both morning and evening quantities will be affected. Thus higher elasticity values are observed. When comparing tables 10 and 11, the elasticities for morning and evening from the demand and full samples are similar in magnitude. Only the all day elasticities seem to differ. The all day sample has much fewer observations than morning and evening samples, so it would be more prone to changes in observation quantity. For the morning and evening samples, it seems that our sub-sample represents the elasticity of the full sample fairly well.

Table 10: Average own-political elasticity summaries by time demand sample

	Morning	**Evening**	**All Day**	**Total**
OLS	0.02521	0.02683	0.03272	0.02645
	(0.01111)	(0.01102)	(0.01121)	(0.01108)
IV	0.02663	0.02834	0.03455	0.02794
	(0.01173)	(0.01164)	(0.01184)	(0.01170)
Nested OLS	0.02534	0.02760	0.03403	0.02705
	(0.01111)	(0.01110)	(0.01102)	(0.01116)
Nested IV	0.02682	0.02882	0.03505	0.02834
	(0.01179	(0.01171)	(0.01177)	(0.01177)
Observations	150	375	7	532

Notes: The standard errors are included in brackets.

Table 11: Average own-political elasticity summaries by time full sample

	Morning	Evening	All Day	Total
OLS	0.02312	0.02597	0.03816	0.02533
	(0.01251)	(0.01202)	(0.01284)	(0.01234)
IV	0.02442	0.02743	0.0403	0.02675
	(0.01321)	(0.01269)	(0.01355)	(0.01303)
Nested OLS	0.02375	0.02731	0.04016	0.02647
	(0.01283)	(0.01238)	(0.01385)	(0.01274)
Nested IV	0.02482	0.02816	0.04138	0.02739
	(0.01340)	(0.01288)	(0.01398)	(0.01324)
Observations	239	566	13	818

Notes: The sample here is the total sample used to calculate market shares. The standard errors are included in brackets.

When separated by time, nested elasticities are twice as large in magnitude as their non-nested counterparts. Even though the nested share coefficient is not significant in the IV specification, the elasticities are still quite different. The standard error for table 11 is much larger than for table 10. Standard errors are around 60 percent of the average elasticity value in table 10 while the standard error of table 11 have similar values to the average elasticities. The number of parameters to be estimated in table 11 is much larger. The largest market in table 10 had 3 newspapers, but the largest market for table 11 is New York City with 26 newspapers. Presence of larger markets creates outliers that may increase the variance of the calculated elasticities.

Furthermore, evening group seem to have higher cross elasticities than the morning group, but our average elasticities are not statistically significant. The difference seems to exist even though the regression nested coefficient gives weak support for the nested structure. The subsample regression analysis in table 7 for nested structure, however, did lend strong support for time based nest structure.

Table 12: Average cross-political elasticity summaries by time for demand sample

	Morning	**Evening**	**Mixed**
OLS	0.0008842	0.002678 (0.001842)	0.003122 (0.002763)
IV	0.0009338	0.002828 (0.001945)	0.003301 (0.002943)
Nested OLS	0.002837	0.005876 (0.0030015)	0.009799 (0.006227)
Nested IV	0.001794	0.004244 (0.002461)	0.006242 (0.004396)
Observations	1	35	117

Notes: Columns 1 and 2 can only be calculated if a given city has multiple newspapers in the same publishing time slot. There was only one city that had two morning edition newspapers. Standard errors are included in brackets.

Table 13: Average cross-political elasticity summaries by time for full sample

	Morning	**Evening**	**Mixed**
OLS	0.0005165 (0.001938)	0.002120 (0.002889)	0.002223 (0.003536)
IV	0.0005454 (0.002047)	0.002239 (0.003051)	0.002221 (0.003578)
Nested OLS	0.0009212 (0.002984)	0.004174 (0.004582)	0.05836 (0.007412)
Nested IV	0.0007255 (0.002510)	0.003150 (0.003794)	0.003860 (0.005255)
Observations	164	175	438

Notes: Same analysis as table 11 but with all 818 papers.

5.3 General Analysis and Extensions

Table 14: Other regression specifications

	OLS full sample	**OLS demand sample**	**OLS (D/R/I)**
Constant	-2.149** (0.1102)	-2.236** (0.1285)	-2.4926** (0.1289)
Subscription price	0.0056 (0.0144)	0.0158 (0.0172)	0.02787 (0.01625)
Political 1			0.2721** (0.05658)
Political 2			0.3258** (0.08114)
R-squared	0.0002	0.0016	0.0406
Observations	785	531	767

Notes: Dependent variable is $\ln(s_j)-\ln(s_o)$. Column 1 contains OLS regression on the full sample using subscription price as the sole regressor. Not all subscription prices were available for the full sample. Column 2 is reproduced from table 1 for comparison. Column 3 contains OLS regression with Democratic, Republican, and Independent as categorical variable for political affiliation. Base group is Democratic, political 1 is Republican, and political 2 is independent. Standard error included in brackets. ** indicates 99% level of significance.

I included full sample in this OLS regression above. Even though I do not have the political affiliation data, I still included those observations in the OLS regression. Compared to original OLS regression where I had mostly monopoly markets, the coefficient here is much closer to zero. In monopoly markets I expect more readership share when the publishers devote more resources into the paper, hire better editors to write more qualities stories. Their higher prices would reflect on the rent charged for better quality. In the regression above I have more competitive markets included in the sample. Here the price competition would counteract the quality component of price. Lower subscription prices will also lead to lower qualities, making the paper less attractive to potential readers.

The independent newspapers are included as an extra specification. The inclusion of independent papers did not change the value of political coefficient between Republican and Democratic, but the R-squared value has drastically increased and the magnitude of the price coefficient also increased. Perhaps the finance and legal papers, ones who rarely declare affiliations, are less prone to government distortions in the form of official purchases. Their price may be more reflective of the paper quality and thus is able to explain the share variations more completely.

In most of my regressions, the explanatory power for the model is very low. A few special circumstances during the 1920s would have affected our data. Government orders can significantly skew the circulation numbers, in particular Wisconsin government had allow each official to order up to 30 copies at no cost of their own (Dyer, 1989). Delivery subsidies provided by the government would also impact the subscription cost.

Instruments for literacy, number of newspaper staff should augment the explaining power of my regression, as they'll help me to better predict the readership base and quality of the paper. A proxy for newspaper quality is the number of pages, which is correlated with market size and higher subscription prices. (Gentzkow, Shapiro, Petek, and Sinkinson, 2015). For recent studies, education, income levels, and urbanization are used as instruments to estimate the demand parameters (Fan, 2012). If data on content devoted to local news, news variety (measured using squared share of staff working in different sections) were available, they could also be included in the regression to support newspaper differentiation (Fan, 2012).

Unfortunately many data sets are proprietary and not freely available for empirical industrial organization research. To make my findings more conclusive, better instruments will be needed to model the demand and the supply side, and even a larger sample that includes more competitive markets maybe required to obtain a negative price coefficient.

6. Conclusion

In this essay, I have estimated a discrete-choice model of demand using various specifications to analyze consumer preferences in the U.S. newspaper market. I found that in accordance with earlier research, consumers do distinguish between Republican and Democratic papers. I detect such distinction in the cross-political elasticity values as well as the political affiliation regression coefficients. The time-of-day based nested logit estimation on large sample did not produce a significant nested logit coefficient, but the coefficient was highly significant in the competitive markets subsample. This suggests that the composition of the sample can be sensitive for arriving at the final results. The cross-product elasticities for morning and evening groups were also different. This time related evidence supports prior belief that it is difficult for newspapers to occupy the same time and political dimensions (Hamilton, 2006). My demand estimation sample generally produced similar elasticity values as the full data sample.

Based on my results, it seems plausible to use the nested logit discrete-choice system to study consumer preferences regarding product characteristics. These studies may help antitrust authorities do define a market that is relevant from consumers' perspectives.

7. References

Ackerberg, Daniel A., and Marc Rysman. 2005. "Unobserved Product Differentiation in Discrete Choice Models: Estimating Price Elasticities and Welfare Effects." *RAND Journal of Economics*. 36(4): 1-19.

Alterman, Eric. 2003. *What Liberal Media? The Truth about Bias and the News.* New York: Basic Books.

Baldasty, Gerald J. 1992 *The Commercialization of News in the Nineteenth Century*. University of Wisconsin Press.

Berry, Steven T. 1994. "Estimating Discrete-Choice Models of Product Differentiation." *RAND Journal of Economics*, 25(2): 242-262.

Berry, Steven, and Panle Jia. 2010. "Tracing the Woes: An Empirical Analysis of the Airline Industry." *American Economics Journal, Microeconomics*, 2(3): 1-43.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. “Automobile Prices in Market Equilibrium.” *Econometrica*, 63(4): 841-890.

Busterna, John C., and Robert G. Picard. 1993. *Joint Operating Agreements: The Newspaper Preservation Act and its Application.* Norwood, N.J.: Ablex Publishing Co.

Chiang, Chun-Fang. 2010. “Political Differentiation in Newspaper Markets.” Unpublished.

Coulter, Ann H. 2003. *Slander: Liberal Lies about the American Right*. New York: Three Rivers Press.

Dubin, J. and D. McFadden. 1984. “An Econometric Analysis of Residential Electric Appliance Holdings and Consumption,” *Econometrica*, 52(2): 345-362.

Dyer, Carolyn S. (1989). *Political Patronage of the Wisconsin Press, 1849–1860: New Perspectives on the Economics of Patronage*. Association for Education in Journalism and Mass Communication.

Fan, Ying. 2012. “Ownership Consolidation and Product Characteristics: A Study of the U.S. Daily Newspaper Market.” 2013. *American Economic Review*. 103(5): 1598-1628.

Gentzkow, Matthrew. “Valuing New Goods in a Model with Complementarity: Online Newspapers.” *American Economic Review*, 97(3): 713-744.

Gentzkow, Matthew, Edward L. Glaeser, and Claudia Goldin (2006). “The Rise of the Fourth Estate: How Newspapers Became Informative and Why it Mattered.” In *Corruption and Reform: Lessons from America’s Economic History*, edited by E. L. Glaeser and C. Goldin. University of Chicago Press, pp. 187–230.

Gentzkow, Matthrew, and Jesse M. Shapiro. 2008. “Competition and Truth in the Market for News.” *Journal of Economic Perspective*, 22(2): 133-154.

Gentzkow, Matthrew, Jesse M. Shapiro, Nathan Petek, and Michael Sinkinson. 2015 “Do Newspapers Serve the State? Incumbent Party Influence on the US Press, 1869-1928.” *Journal of the European Economic Association*, 13(1): 29-61.

Gentzkow, Matthrew, Jesse M. Shapiro, and Michael Sinkinson. 2014. “Competition and Ideological Diversity: Historical Evidence from US Newspapers.” *American Economic Review*, 104(10): 3073-3114.

Gentzkow, Matthrew, Jesse M. Shapiro, and Michael Sinkinson. 2011. “The Effect of Newspaper Entry and Exit on Electoral Politics.” *American Economics Review*, 101(7): 2980-3018.

Goldberg, Bernard. 2003. *Bias: A CBS Insider Exposes How the Media Distort the News*. Washington, DC: Regnery.

Graber, Doris A, ed. 2000. *Media Power in Politics.* Washington, DC: CQ Press College.

Greene, William H. ed. 2011. Econometric Analysis. Toronto: Pearson Education Limited.

Hamilton, James T. 2006. *All the News That's Fit to Sell: How the Market Transforms Information into News*. Princeton: Princeton University Press.

Hendel, Igal. 1999. "Estimating Multiple-Discrete Choice Models: An Appliction to Computerization Returns." *The Review of Economic Studies*, 66(2): 423-446.

Lord, Charles G., Lee Ross, and Mark R. Lepper. 1979. "Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence." *Journal of Personality and Social Psychology*, 37(11): 2098–2109.

Mazur, Joe. 2011. "Sensitivity Analysis in Merger Simulation." *Working Paper*. Duke University, Durham.

Mondak, Jeffery J. 1995. "Newspapers and Political Awareness." *American Journal of Political Science,* 39(2): 513–27.

Nisbett, Richard, and Lee Ross. 1980. *Human Inference: Strategies and Shortcomings of Social Judgment*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Peters, Craig. 2006. "Evaluating the Performance of Merger Simulation: Evidence from the U.S. Airline Industry." *Journal of Law and Economics*, 49(2): 627-649.

Prat, Andrea, and David Stromberg. 2005. "Commercial Television and Voter Information." CEPR Discussion Papers 4989.

Snyder, James M., Jr., and David Strömberg. 2008. "Press Coverage and Political Accountability." National Bureau of Economic Research Working Paper 13878.

Stewart, James B. 1999. "Consider the Sources." *New York Times,* July 4.

Strömberg, David. 2004. "Radio's Impact on Public Spending." *Quarterly Journal of Economics*, 119(1): 189–221.

Zaller, John. 1999. "Market Competition and News Quality." Paper presented at the

1999 Annual Meetings of the American Political Science Association, Atlanta, GA. http:// www.sscnet.ucla.edu/polisci/faculty/zaller/ News%20quality%20paper.PDF.

8. Appendix

8.1 Derivation of nested logit elasticity

Nested logit model: own-political elasticity

$$f_j(\delta,\sigma) = f_{j|g}(\delta,\sigma) f_g(\delta,\sigma) = \frac{e^{\delta_j/1-\sigma}}{D_g^{\sigma}[\sum_g D_g^{1-\sigma}]}$$

$$\text{where } D_g = \sum_{j\in g} e^{\delta_j/1-\sigma}$$

$$\frac{\partial f_j}{\partial x_j}$$

$$= \frac{D_g^{\sigma}[\sum_g D_g^{1-\sigma}] e^{\delta_j/1-\sigma}\left(\frac{\beta}{1-\sigma}\right) - e^{\delta_j/1-\sigma}\left[\sigma D_g^{\sigma-1} e^{\delta_j/1-\sigma}\left(\frac{\beta}{1-\sigma}\right)\sum_g D_g^{1-\sigma} + D_g^{\sigma} D_g^{-\sigma}(1-\sigma) e^{\delta_j/1-\sigma}\left(\frac{\beta}{1-\sigma}\right)\right]}{D_g^{2\sigma}\left[\sum_g D_g^{1-\sigma}\right]^2}$$

$$= \frac{e^{\delta_j/1-\sigma}\left(\frac{\beta}{1-\sigma}\right)[D_g^{\sigma}[\sum_g D_g^{1-\sigma}] - e^{\delta_j/1-\sigma}\left[\sigma D_g^{\sigma-1}\sum_g D_g^{1-\sigma} + 1 - \sigma\right]]}{(D_g^{\sigma}[\sum_g D_g^{1-\sigma}])^2}$$

$$= \frac{f_j\beta\frac{1}{1-\sigma}[D_g^{\sigma}[\sum_g D_g^{1-\sigma}] - e^{\delta_j/1-\sigma}\left[\sigma D_g^{\sigma-1}\sum_g D_g^{1-\sigma} + 1 - \sigma\right]]}{D_g^{\sigma}[\sum_g D_g^{1-\sigma}]}$$

$$= f_j\beta\left(\frac{1}{1-\sigma}\left[1 - \sigma e^{\delta_j/1-\sigma} D_g^{-1} + (1-\sigma)\frac{e^{\delta_j/1-\sigma}}{D_g^{\sigma}\left[\sum_g D_g^{1-\sigma}\right]}\right]\right)$$

$$= f_j\beta\left(\frac{1}{1-\sigma} - \frac{\sigma}{1-\sigma} f_{j|g} + f_j\right)$$

where $f_{j|g}(\delta,\sigma) = \frac{e^{\delta_j/1-\sigma}}{D_g}$

Nested logit model: cross-political elasticity. Goods j and k belong to the same nest.

$$\frac{\partial f_j}{\partial x_k} = -\frac{0 - e^{\delta_j/_{1-\sigma}}\left[\sigma D_g^{\sigma-1} e^{\delta_k/_{1-\sigma}}\left(\frac{\beta}{1-\sigma}\right)\sum_g D_g^{1-\sigma} + D_g^{\sigma} D_g^{-\sigma}(1-\sigma)e^{\delta_k/_{1-\sigma}}\left(\frac{\beta}{1-\sigma}\right)\right]}{D_g^{2\sigma}\left[\sum_g D_g^{1-\sigma}\right]^2}$$

$$= \frac{e^{\delta_j/_{1-\sigma}}\left(\frac{\beta}{1-\sigma}\right)e^{\delta_k/_{1-\sigma}}\left[\sigma D_g^{\sigma-1}\left[\sum_g D_g^{1-\sigma}\right] + 1 - \sigma\right]}{\left(D_g^{\sigma}\left[\sum_g D_g^{1-\sigma}\right]\right)^2}$$

$$= \frac{f_k\beta \frac{1}{1-\sigma} e^{\delta_j/_{1-\sigma}}\left[\sigma D_g^{\sigma-1}\left[\sum_g D_g^{1-\sigma}\right] + 1 - \sigma\right]}{D_g^{\sigma}\left[\sum_g D_g^{1-\sigma}\right]}$$

$$= \frac{f_k\beta\left[e^{\delta_j/_{1-\sigma}}\frac{\sigma}{1-\sigma}D_g^{\sigma-1}\left[\sum_g D_g^{1-\sigma}\right] + e^{\delta_j/_{1-\sigma}}\right]}{D_g^{\sigma}\left[\sum_g D_g^{1-\sigma}\right]}$$

$$= f_k\beta\left(\frac{\sigma}{1-\sigma}f_{j|g} + f_j\right)$$

Nested logit model: cross-political elasticity. Goods j and k belong to the different nests, nests g and j.

$$\frac{\partial f_j}{\partial x_k} = -\frac{0 - e^{\delta_j/_{1-\sigma}}\left[0 + D_g^{\sigma} D_j^{-\sigma}(1-\sigma)e^{\delta_k/_{1-\sigma}}\left(\frac{\beta}{1-\sigma}\right)\right]}{D_g^{2\sigma}\left[\sum_g D_g^{1-\sigma}\right]^2}$$

$$= \frac{e^{\delta_j/_{1-\sigma}}\left(\frac{\beta}{1-\sigma}\right)e^{\delta_k/_{1-\sigma}}(1-\sigma)D_g^{\sigma}D_j^{-\sigma}}{\left(D_g^{\sigma}\left[\sum_g D_g^{1-\sigma}\right]\right)^2}$$

$$= \frac{e^{\delta_j/_{1-\sigma}}\left(\frac{\beta}{1-\sigma}\right)e^{\delta_k/_{1-\sigma}}(1-\sigma)D_j^{-\sigma}}{\left(D_g^{\sigma}\left[\sum_g D_g^{1-\sigma}\right]\right)\left(\left[\sum_g D_g^{1-\sigma}\right]\right)}$$

$$= \frac{e^{\delta_j/_{1-\sigma}}\beta e^{\delta_k/_{1-\sigma}}}{\left(D_g^{\sigma}\left[\sum_g D_g^{1-\sigma}\right]\right)\left(D_j^{\sigma}\left[\sum_g D_g^{1-\sigma}\right]\right)} = \beta f_k f_j$$

8.2 Fixed cost IV

Table 15: IV regression including fixed cost

Demand variables	IV	IV first stage	IV	IV first stage
Political affiliation	0.2931**		0.2890**	
	(0.06388)		(0.0626)	
Constant	-3.813**		-3.601**	
	(0.6213)		(0.6373)	
Price	0.2063*		0.1777*	
	(0.08380)		(0.0859)	
Political		-0.06088		-0.0668

affiliation		(0.1446)		(0.1442)
Advertising		-0.0327		-0.0408
revenue		(0.0283)		(0.0246)
Variable cost		0.1253**		0.1255**
		(0.0422)		0.0425
Political competitors		0.3443*		0.3413*
		(0.1541)		(0.1539)
Fixed cost		-0.01331		
		(0.02341)		
Constant		6.291**		6.311**
		(0.3219)		(0.3201)
R-squared	0.0397		0.0391	
Adjusted R-squared	0.0305		0.0317	
Observations	528		528	

Notes: The sample only includes the observations with Republican or Democratic as declared political affiliation. Dependent variable is ln(s_j)-ln(s_o). First stage results for 2SLS are also reported. Robust standard errors are included brackets. * indicates 95% level of significance. ** indicates 99% level of significance

8.3 Stata Code (For Online Version Only)

```
//  ****----- INITIAL VARIABLE GENERATION
set more off
clear all
use "/Volumes/LEXAR YX/Important stuff/useful data/basic
info/combined_set2.dta"
//Two lines of data cleaning
//drop duplicate avgDAPC
//drop if _merge == 2

//generating share, outshares, explained variable
generate share = DailyAverageNetPaidCirculation/PopulationOfTradingTerritory
egen inshare = total(share), by(City State)
gen outshare = 1-inshare
gen lndiff = ln(share)-ln(outshare)
rename polaff polaff1
//merge to get another set of polaff
drop _merge

//getting nested regressor
gen ln_nest_time_share = ln(share/total_time_share)

sort City State time
by City State time: gen time_comp=_N
gen pol_aff = 1 if polaff1=="R"
replace pol_aff = . if pol_aff!=1
replace pol_aff = 0 if polaff1=="D"


// ****----- REGRESSION AND ELASTICITY CALCULATIONS
set more off
clear all
//use "F:\Important stuff\useful data\basic info\combined_set3s.dta"
use "/Volumes/LEXAR YX/Important stuff/useful data/basic
info/combined_set3m.dta"

//keeping only observations with positive shares
keep if DailyAverageNetPaidCirculation!=.
drop if share==.

//getting number of competitors in each market
sort City State time
quietly by City State time: gen time_comp = _N
drop if missing(pol_aff)
//regular and iv regression on depend var lnsj-lns_out
```

```
//pol_aff 0=Democrat 1=Repub

// ----*** logit regression ***---
regress lndiff pol_aff subprice, vce(robust)
generate ols_logit_own_elas=_b[pol_aff]*share*(1-share)
// drop if pol_aff==.
sort City State time NewspaperName

//series of cross elasticities

gsort City State -time NewspaperName
local count=1
while `count'<3 {
	quietly by City State: generate
morn_ols_logit_cross_elas_`count'=_b[pol_aff]*share*share[`count'] if
time[`count']=="morn" & time=="morn" & _n>`count'
	local count = `count'+1
}
sort City State time NewspaperName
local count=1
while `count'<3 {
	quietly by City State: generate
even_ols_logit_cross_elas_`count'=_b[pol_aff]*share*share[`count'] if
time[`count']=="eveng" & time=="eveng" & _n>`count'
	local count = `count'+1
}
sort City State time NewspaperName
local count=1
while `count'<3 {
	quietly by City State: generate
mix_ols_logit_cross_elas_`count'=_b[pol_aff]*share*share[`count'] if
time[`count']!=time & _n>`count'
	local count = `count'+1
}

// ** Cross elasticities ended **
sort City State pol_aff
order morn_ols_logit_cross_elas_* mix_ols_logit_cross_elas_* City State time
even_ols_logit_cross_elas_*, after(Recid)



ivregress 2sls lndiff pol_aff (subprice=adv_rev_per_copy variable_cost_per_copy
ideo_comp), first vce(robust)
//political elasticities
generate iv_logit_own_elas=_b[pol_aff]*share*(1-share)
```

```
//series of cross elasticities
gsort City State -time NewspaperName
local count=1
while `count'<3 {
	quietly by City State: generate
morn_iv_logit_cross_elas_`count'=_b[pol_aff]*share*share[`count'] if
time[`count']=="morn" & time=="morn" & _n>`count'
	local count = `count'+1
}
sort City State time NewspaperName
local count=1
while `count'<3 {
	quietly by City State: generate
even_iv_logit_cross_elas_`count'=_b[pol_aff]*share*share[`count'] if
time[`count']=="eveng" &time=="eveng" & _n>`count'
	local count = `count'+1
}
local count=1
while `count'<3 {
	quietly by City State: generate
mix_iv_logit_cross_elas_`count'=_b[pol_aff]*share*share[`count'] if
time[`count']!=time & _n>`count' & time!="AD" & time[`count']!="AD"
	local count = `count'+1
}
order morn_ols_logit_cross_elas_* morn_iv_logit_cross_elas_*
mix_ols_logit_cross_elas_* mix_iv_logit_cross_elas_* City State pol_aff
even_ols_logit_cross_elas_* even_iv_logit_cross_elas_*, after(Recid)

// ***---- NESTED LOGIT ------****

gen ln_nest_time_share=ln(nest_time_share)
regress lndiff pol_aff subprice ln_nest_time_share, vce(robust)
generate ols_nest_own_elas=_b[pol_aff]*share*(1/(1-_b[ln_nest_time_share])-
_b[ln_nest_time_share]*nest_time_share/(1-_b[ln_nest_time_share])-share)

//Series of cross elasticities
gsort City State -time NewspaperName
local count=1
while `count'<3 {
		quietly by City State: generate
morn_ols_nest_cross_elas_`count'=_b[pol_aff]*share[`count']*(_b[ln_nest_time_share
]*nest_time_share/(1-_b[ln_nest_time_share])+share) if time[`count']=="morn" &
time=="morn" & _n>`count'
//		quietly by City State: replace
morn_ols_nest_cross_elas_`count'=_b[pol_aff]*share*share[`count'] if
time[`count']=="morn" & time=="morn" & _n!=`count' & time!=time[`count']
```

```
        local count = `count'+1
}
sort City State time NewspaperName
local count=1
while `count'<3 {
                quietly by City State: generate
even_ols_nest_cross_elas_`count'=_b[pol_aff]*share[`count']*(_b[ln_nest_time_share]
*nest_time_share/(1-_b[ln_nest_time_share])+share) if time[`count']=="eveng" &
time=="eveng" & _n>`count' //& time==time[`count']
        local count = `count'+1
}
local count=1
while `count'<3 {
                quietly by City State: generate
mix_ols_nest_cross_elas_`count'=_b[pol_aff]*share[`count']*(_b[ln_nest_time_share]*
nest_time_share/(1-_b[ln_nest_time_share])+share) if time!="AD" &
time[`count']!="AD" & _n>`count' & time[`count']!=time //& time==time[`count']
        local count = `count'+1
}
// OLS cross elasticity ended

order morn_ols_nest_cross_elas_* mix_ols_nest_cross_elas_* share nest_time_share
City State time even_ols_nest_cross_elas_*, after(Recid)


// **** ----- NESTED LOGIT IV -------****
ivregress 2sls lndiff pol_aff ln_nest_time_share (subprice=adv_rev_per_copy
variable_cost_per_copy ideo_comp), first vce(robust)
generate iv_nest_own_elas=_b[pol_aff]*share*(1/(1-_b[ln_nest_time_share])-
_b[ln_nest_time_share]*nest_time_share/(1-_b[ln_nest_time_share])-share)
sort City State time NewspaperName

//Series of cross elasticities
gsort City State -time NewspaperName
local count=1
while `count'<3 {
                quietly by City State: generate
morn_iv_nest_cross_elas_`count'=_b[pol_aff]*share[`count']*(_b[ln_nest_time_share]*
nest_time_share/(1-_b[ln_nest_time_share])+share) if time[`count']=="morn" &
time=="morn" & _n>`count' //& time==time[`count']
                local count = `count'+1
}
sort City State time NewspaperName
local count=1
while `count'<3 {
                quietly by City State: generate
```

```
even_iv_nest_cross_elas_`count'=_b[pol_aff]*share[`count']*(_b[ln_nest_time_share]*
nest_time_share/(1-_b[ln_nest_time_share])+share) if time[`count']=="eveng" &
time=="eveng" & _n>`count' //& time==time[`count']
		local count = `count'+1
}
local count=1
while `count'<3 {
		quietly by City State: generate
mix_iv_nest_cross_elas_`count'=_b[pol_aff]*share[`count']*(_b[ln_nest_time_share]*n
est_time_share/(1-_b[ln_nest_time_share])+share) if time[`count']!=time &
_n>`count' & time!="AD" &time[`count']!="AD"
		local count = `count'+1
}
// IV cross elasticity ended
save "/Volumes/LEXAR YX/Important stuff/useful data/basic
info/elasticities_bytime_samll_Aug24.dta"

set more off
clear all
use "/Volumes/LEXAR YX/Important stuff/useful data/basic
info/elasticities_bytime_samll_Aug24.dta"

// stacking the elasticities to summarize them
stack morn_ols_nest_cross_elas_*, into(morn_ols_nest_cross_elas) clear wide
sum morn_ols_nest_cross_elas

// ****----Different specifications
set more off
clear all
//use "F:\Important stuff\useful data\basic info\combined_set3s.dta"
use "/Volumes/LEXAR YX/Important stuff/useful data/basic
info/combined_set3m.dta"
keep if DailyAverageNetPaidCirculation!=.
drop _merge
merge 1:1 Recid using "/Volumes/LEXAR YX/Important stuff/useful data/basic
info/ucab_yearly_inland_combined2_polaff.dta"

drop if share==.
gen polaff_incomp=polaff_num
replace polaff_num=2 if polaff_num==. & polaff1=="I"
drop _merge
rename subprice subprice_orig
merge 1:1 Recid using "/Volumes/LEXAR YX/Important stuff/useful data/basic
info/ucab_yearly_inland_combined2_subprice.dta"
drop if share==.
sort City State time
```

```
quietly by City State time: gen time_comp = _N
//regular and iv regression on depend var lnsj-lns_out
//pol_aff 0=Democrat 1=Repub
//Categorical regression for table 14
regress lndiff i.polaff_num subprice, vce(robust)


//appendix regression having fixed cost as IV
ivregress 2sls lndiff pol_aff (subprice=adv_rev_per_copy variable_cost_per_copy
ideo_comp), first vce(robust)
ivregress 2sls lndiff pol_aff (subprice=adv_rev_per_copy fixed_cost_per_copy
variable_cost_per_copy ideo_comp), first vce(robust)

//Looking at competitive markets for time edition differentiation
gen ln_nest_time_share=ln(nest_time_share)
sort City State time
by City State time: gen ed_comp = _N
ivregress 2sls lndiff pol_aff ln_nest_time_share (subprice=adv_rev_per_copy
variable_cost_per_copy ideo_comp) if ed_comp>1, first vce(robust)
```