where $g_s(\mathcal{B})$ is the number of distinct two-newspaper subsets of bundle $\mathcal{B}$ such that the two newspapers have the same political affiliation, $g_d(\mathcal{B})$ is the number of two-newspaper subsets with different affiliations, and $\varepsilon_{im}(\mathcal{B})$ is a type-I extreme value error. Note that the utility from consuming no newspapers is $\varepsilon_{im}(\emptyset)$. A household thus receives per-newspaper utility $\overline{\beta}$ for each newspaper in its consumption bundle that has the same affiliation as the household, and per-newspaper utility $\underline{\beta}$ for each newspaper that has a different affiliation. Utility is diminished by an amount Γ_s for every pair of newspapers with the same affiliation and by an amount Γ_d for every pair with a different affiliation. The specification thus allows that same-affiliation papers are closer substitutes than opposite-affiliation papers. We assume that this demand specification applies to both newspaper markets and hinterland towns.

C. *Equilibrium*

We derive a pure-strategy perfect Bayesian equilibrium of the model beginning at the end of the game and working backward.

In the final stage of the game, the demands q_{jm} are uniquely determined given the number of newspapers, their affiliations, and their circulation prices. Integration over ε_{im} and θ_{im} yields a closed form for q_{jm} as a sum of familiar logit probabilities.[20]

In the fourth stage, newspapers simultaneously choose advertising prices given the number of newspapers, their affiliations, and their circulation prices. In any pure strategy equilibrium of the advertising pricing stage in market m with affiliations $\boldsymbol{\tau}$ and prices $\mathbf{p}$, all advertisers must advertise in all newspapers ($\psi_{jm} = 1$), and newspaper j's advertising price per copy must equal

$$a_{jm}(\mathbf{p}, \boldsymbol{\tau}) = a_h \mathcal{E}_{jm}(\mathbf{p}, \boldsymbol{\tau}) + a_l\big(1 - \mathcal{E}_{jm}(\mathbf{p}, \boldsymbol{\tau})\big), \tag{6}$$

where $\mathcal{E}_{jm}$ is the share of newspaper j's readers who are "exclusive" in the sense that they read no other newspaper.[21] In equilibrium, each newspaper charges advertisers

[20] Let

$$u_m^\theta(\mathcal{B}) = \sum_{j \in \mathcal{B}} \left(\underline{\beta}\mathbf{1}_{\theta \neq \tau_{jm}} + \overline{\beta}\,\mathbf{1}_{\theta = \tau_{jm}} - \alpha p_{jm}\right) - g_s(\mathcal{B})\Gamma_s - g_d(\mathcal{B})\Gamma_d \tag{3}$$

denote the mean utility of households of type θ for bundle $\mathcal{B}$. Then the share of households of type θ who purchase newspaper j is

$$q_{jm}^\theta = \frac{\sum_{\{\mathcal{B} \in \mathbb{B} \,:\, j \in \mathcal{B}\}} \exp\big(u_m^\theta(\mathcal{B})\big)}{\sum_{\mathcal{B}' \in \mathbb{B}} \exp\big(u_m^\theta(\mathcal{B}')\big)}, \tag{4}$$

where $\mathbb{B}$ is the set of all bundles of the papers in market m. The market-wide share of households purchasing newspaper j is then

$$q_{jm} = \rho_m q_{jm}^R + (1 - \rho_m)\, q_{jm}^D. \tag{5}$$

[21] Although demand has not yet been realized at the advertising stage, $\mathcal{E}_{jm}$ depends only on affiliations and prices, both of which have been chosen at this stage of the game. Anderson, Foros, and Kind (2011) prove our characterization formally. A proof sketch is as follows. First, observe that in any equilibrium all advertisers must advertise in all newspapers, since if a newspaper receives no advertising, there is always some positive advertising price below a_l that the newspaper would like to charge and that would attract advertising, thus raising the newspaper's profits. Second, observe that in any equilibrium each newspaper will charge a price such that advertisers are indifferent between advertising in that newspaper and not; otherwise the newspaper could raise its advertising price and increase its profits. With all advertisers advertising in all newspapers, it is straightforward to show that this maximum price is given by (6) for all newspapers, implying the desired result.

only for the incremental value of the impressions the newspaper can deliver, which is reduced if these impressions are duplicated with other newspapers.

In the third stage, firms simultaneously choose circulation prices given the number of newspapers and their affiliations. An equilibrium of this stage in market m with affiliations $\boldsymbol{\tau}$ is a vector $\mathbf{p}^*$ such that each element p_j^* satisfies

$$p_j^* \in \arg\max_{p_j} \left(p_j + a_{jm}(\mathbf{p}^*, \boldsymbol{\tau}) - MC\right) q_{jm}(\mathbf{p}^*, \boldsymbol{\tau}). \tag{7}$$

We cannot provide a proof of the uniqueness of the pricing game equilibrium. In estimation we solve numerically for the first-order conditions of the game and we verify that all newspapers' second-order conditions hold at the solution. We choose a starting value close to the observed prices and verify that the solution is not sensitive to local variation (plus or minus \$1 per copy) in the choice of starting value at the estimated parameters.

In the second stage, firms sequentially choose affiliations given the number of newspapers and their affiliation-specific shocks ξ_{jm}. An equilibrium of this stage in market m given the number of newspapers J is a vector $\boldsymbol{\tau}^*$ such that each τ_j^* maximizes $\left(\mathrm{E}_{\boldsymbol{\tau}_{j^+}^*} v_{jm}\left(\left[\boldsymbol{\tau}_{j^-}^*, \tau_j^*, \boldsymbol{\tau}_{j^+}^*\right]\right) - \xi_{jm}(\tau_j^*)\right)$, where $\boldsymbol{\tau}_{j^-}^*$ and $\boldsymbol{\tau}_{j^+}^*$ are vectors of affiliations of the newspapers with indices less than and greater than j, respectively, and $v_{jm}(\boldsymbol{\tau})$ denotes the equilibrium value of $(p_{jm} + a_{jm} - MC)\, q_{jm}$ given affiliations $\boldsymbol{\tau}$. For generic realizations of cost shocks ξ_{jm} there is a unique equilibrium vector of affiliation choices that can be computed by backward induction.

In the first stage, potential entrants sequentially choose either to enter or to not enter. At this point in the game all potential entrants are symmetric and share the same information sets, and since the number of potential entrants is finite, this stage has a unique equilibrium for generic parameter values. Let $P_m(\boldsymbol{\tau})$ denote the equilibrium probability that the second-stage affiliation vector is $\boldsymbol{\tau}$ conditional on $|\boldsymbol{\tau}|$ newspapers entering. Given affiliations $\boldsymbol{\tau}$, let $\bar{\xi}_{jm}(\boldsymbol{\tau})$ denote the expected value of $\xi_{jm}(\tau_j)$ conditional on newspaper j choosing its affiliation optimally. The per-household expected variable profit of each entering newspaper given J entrants is

$$V_m(J) = \frac{1}{J} \sum_{j=1}^{J} \sum_{\boldsymbol{\tau} \in \mathcal{T}_J} \left(v_{jm}(\boldsymbol{\tau}) - \bar{\xi}_{jm}(\boldsymbol{\tau})\right) P_m(\boldsymbol{\tau}), \tag{8}$$

where $\mathcal{T}_J$ is the set of $\boldsymbol{\tau}$ vectors with $|\boldsymbol{\tau}| = J$. If V_m is strictly decreasing in J, the equilibrium number of firms J^* is the unique number such that entering newspapers are profitable but a marginal entrant would not be. That is,

$$V_m(J^*) \geq \frac{\kappa_m}{S_m} > V_m(J^* + 1) \tag{9}$$

for $J^* \in \{1, \ldots, J^{max} - 1\}$. If $V_m(1) < \kappa_m / S_m$ then $J^* = 0$ is the equilibrium, and if $V_m(J^{max}) > \kappa_m / S_m$ then $J^* = J^{max}$ is the equilibrium. Though we do not have a formal proof that V_m must be decreasing, this condition is intuitive: it means that all else equal a firm would rather be in a market with fewer competitors. In repeated simulations we find that this property holds for all markets in our data at the estimated parameters.

D. *Discussion*

Market Definition.—We make two important simplifying assumptions in defining newspaper markets. First, we assume that newspapers only compete with other newspapers headquartered in the same market, and we ignore circulation in hinterland towns in modeling newspapers' affiliation, pricing, and entry choices. In reality, the 1924 ABC data show that home-market papers constitute 90 percent of circulation in news markets, and the average newspaper sold 65 percent of copies in its home market.[22] Our definition is thus an approximation to a reality in which consumers exhibit strong but not exclusive preferences for local papers. To assess robustness to this assumption, we show results in Appendix A from a subsample that excludes markets close to large cities, and from a specification in which we incorporate hinterland towns into our measure of market ideology.

Second, we aggregate all substitutes for daily newspapers into an outside option whose prices and characteristics we do not model explicitly. We deliberately choose a period of study in which there were few such substitutes that were also significant sources of political information. In 1924, television did not exist and radio was in its infancy as a news source (Sterling and Kittross 2001). Although weekly newspapers and magazines existed and played an important role in the media market, neither conveyed the news on a daily basis, and neither weekly newspapers nor weekly magazines achieved total weekly circulation in excess of the total *daily* circulation of daily newspapers (Field 2006).

Product Characteristics.—Our model endogenizes political affiliation but not other forms of differentiation. This is clearly a dramatic simplification, as variation in both quality and nonpolitical horizontal dimensions (such as time of publication) was clearly important. Estimating consumer preferences for other dimensions is straightforward, but endogenizing newspapers' choices of attributes along multiple dimensions would add significant complexity.

Failing to account for unobserved vertical differentiation can lead to bias in price coefficients (Berry, Levinsohn, and Pakes 1995). As we detail further below, this concern motivates us to identify the price coefficient α from the monopoly first-order condition (Gentzkow 2007) rather than from price variation. We present several additional sensitivity analyses. In Appendix A we show results from a model which allows utility to depend on distance to a newspaper's headquarters, an important shifter of quality. We also experiment with specifications which use a newspaper's price and home market circulation as quality proxies. In the online Appendix, we show explicitly that the crucial cross-sectional patterns which identify our demand system are robust to allowing flexibly for variation in quality at the newspaper level.

Ignoring nonpolitical horizontal differentiation is also an important simplification. The more newspapers can differentiate on nonpolitical dimensions, the weaker will be their incentive to differentiate on politics (Irmen and Thisse 1998; Liu and Shuai

[22] Fan (2013) uses a wider market definition which encapsulates 85 percent of a newspaper's circulation and includes circulation outside a newspaper's home market. Allowing for this additional realism would be difficult in our model as we estimate newspapers' incentive to enter the market. Doing so in a model with overlapping markets would mean computing post-entry equilibrium configurations taking account of strategic linkages of distant newspapers through chains of partially overlapping markets.

2013). We expect our empirical estimates to reflect the incentive to differentiate on political dimensions *given* the extent of differentiation on other dimensions. As we do not treat differentiation on other dimensions as endogenous, we cannot allow it to vary in our counterfactuals, which could impact our conclusions. For example, we could overstate or understate the welfare effects of collusion, because collusion can both encourage and discourage nonpolitical differentiation (Sweeting 2010). In addition, we could overstate the value of variety, because we capture nonpolitical differentiation through a symmetric logit error (Ackerberg and Rysman 2005).

In the online Appendix we present sensitivity analyses related to these concerns. We show that our finding that variety is undersupplied survives even if we exogenously cap the number of entering newspapers. We also present results from an experiment in which we simulate data from a model with two types of horizontal differentiation—politics and time of publication—and estimate a misspecified model that allows only for political differentiation. We find that the estimated model matches the qualitative counterfactual predictions of the data-generating model well, although there are some quantitative differences between the two.

Consumer Preferences.—Our demand specification is designed to capture two key elements: consumers' preferences for like-minded political news, and the possibility of a given household reading multiple papers. The former is obviously important given our focus on political differentiation. The latter is equally crucial, because audience duplication across papers will be the key driver of advertising competition. It is also consistent with our readership surveys, which show a significant amount of multiple readership.

Our demand model nests several cases of interest. When $\underline{\beta} = \overline{\beta}$ and $\Gamma_d = \Gamma_s = 0$, it is equivalent to a model in which each newspaper is a monopolist facing logit demand. When $\underline{\beta} = \overline{\beta}$ and $\Gamma_d, \Gamma_s \to \infty$, it is a standard logit model in which each household reads at most one newspaper. When $\underline{\beta} \to -\infty$, it is equivalent to a model in which there are two distinct markets, one for R newspapers and one for D newspapers.

An important simplifying assumption is that we allow only two types of consumers: Republican and Democratic. In reality, of course, some consumers in the period we study did not have a definite partisan affiliation, although this group was likely in a minority.[23] In the presence of unmodeled nonpartisan consumers, we expect that the gap $\overline{\beta} - \underline{\beta}$ which we estimate will measure the "average" level of partisanship in news preferences (i.e., a value between the preferences of partisans and nonpartisans). The online Appendix presents estimates of an augmented demand system which allows for politically unaffiliated consumers, and shows how our counterfactual estimates change with the fraction unaffiliated.

We assume that the same demand model applies in newspapers' headquarter markets and in the small, hinterland towns surrounding those markets. The assumption that preferences do not depend directly on market size, though common in the

[23] Using Burnham's (1965) index based on aggregate election returns, Rusk (1970) estimates that split-ticket voting in the United States was not more than 7 percent during the period 1876–1908. Millspaugh (1918) reports based on actual ballot records that 21.7 percent of votes in Rhode Island in 1906 were split-ticket. Erikson and Tedin (1981) report that 26 percent of voters switched parties between the 1924 and 1928 elections and that this figure fell to 11 percent by 1944. (The 1924–1928 period coincides with a major shift in US party politics.)