Econometrica, Vol. 78, No. 1 (January, 2010), 35–71

WHAT DRIVES MEDIA SLANT? EVIDENCE FROM U.S. DAILY NEWSPAPERS

BY MATTHEW GENTZKOW AND JESSE M. SHAPIRO[1]

We construct a new index of media slant that measures the similarity of a news outlet's language to that of a congressional Republican or Democrat. We estimate a model of newspaper demand that incorporates slant explicitly, estimate the slant that would be chosen if newspapers independently maximized their own profits, and compare these profit-maximizing points with firms' actual choices. We find that readers have an economically significant preference for like-minded news. Firms respond strongly to consumer preferences, which account for roughly 20 percent of the variation in measured slant in our sample. By contrast, the identity of a newspaper's owner explains far less of the variation in slant.

KEYWORDS: Bias, text categorization, media ownership.

1. INTRODUCTION

GOVERNMENT REGULATION OF NEWS MEDIA ownership in the United States is built on two propositions. The first is that news content has a powerful impact on politics, with ideologically diverse content producing socially desirable outcomes. According to the U.S. Supreme Court (1945), "One of the most vital of all general interests [is] the dissemination of news from as many different sources, and with as many different facets and colors as is possible. That interest . . . presupposes that right conclusions are more likely to be gathered out of a multitude of tongues, than through any kind of authoritative selection."

The second proposition is that unregulated markets will tend to produce too little ideological diversity. The highly influential Hutchins Commission report identified cross-market consolidation in newspaper ownership as a major obstacle to the emergence of truth in the press (Commission on Freedom of

[1]We are grateful to Attila Ambrus, David Autor, Gary Becker, Gary Chamberlain, Raj Chetty, Tim Conley, Liran Einav, Edward Glaeser, Tim Groseclose, Christian Hansen, Justine Hastings, Chris Hayes, Daniel Hojman, Matt Kahn, Larry Katz, John List, Kevin M. Murphy, Ben Olken, Ariel Pakes, Andrea Prat, Riccardo Puglisi, Sam Schulhofer-Wohl, Andrei Shleifer, Monica Singhal, Jim Snyder, Wing Suen, Catherine Thomas, Abe Wickelgren, and numerous seminar and conference participants for helpful comments. We especially wish to thank Renata Voccia, Paul Wilt, Todd Fegan, and the rest of the staff at ProQuest for their support and assistance at all stages of this project. Mike Abito, Steve Cicala, Hays Golden, James Mahon, Jennifer Paniza, and Mike Sinkinson provided outstanding research assistance and showed tireless dedication to this project. We also thank Yujing Chen, Alex Fogel, Lisa Furchtgott, Ingrid Gonçalves, Hayden Haralson Hudson, and Hannah Melnicoe for excellent research assistance. This research was supported by National Science Foundation Grant SES-0617658, as well as the Stigler Center for the Study of the State and the Economy, the Initiative on Global Markets, and the Centel Foundation/Robert P. Reuss Faculty Research Fund, all at the University of Chicago Booth School of Business.

 DOI: 10.3982/ECTA7195

the Press (1947)). The Federal Communications Commission (FCC) "has traditionally assumed that there is a positive correlation between viewpoints expressed and ownership of an outlet. The Commission has sought, therefore, to diffuse ownership of media outlets among multiple firms in order to diversify the viewpoints available to the public" (FCC (2003)). This belief has justified significant controls on cross-market consolidation in broadcast media ownership, on foreign ownership of media, and on cross-media ownership within markets, and has motivated a sizable academic literature arguing that current media ownership is too concentrated (Bagdikian (2000)).

That news content can have significant effects on political attitudes and outcomes has been documented empirically by Strömberg (2004), Gentzkow and Shapiro (2004), Gentzkow (2006), Gerber, Karlan, and Bergan (2009), DellaVigna and Kaplan (2007), and others. In contrast, evidence on the incentives that shape ideological content and on the role of ownership, in particular, is limited. Existing studies have generally relied on hand collection and coding of news content, and so have been restricted to small numbers of sources (e.g., Glasser, Allen, and Blanks (1989), Pritchard (2002)). Groseclose and Milyo (2005) made an important contribution, proposing a new measure of ideological content based on counts of think-tank citations. However, their index was calculated only for a small number of outlets, and has not been used to analyze the determinants of slant.

In this paper, we propose a new index of ideological slant in news coverage and compute it for a large sample of U.S. daily newspapers. We estimate a model of newspaper demand that incorporates slant explicitly, estimate the slant that would be chosen if newspapers independently maximized their own profits, and compare these profit-maximizing points with firms' actual choices. We estimate the contributions of consumer and owner heterogeneity to cross-market diversity in slant and develop tentative implications for ownership regulation.

Our slant index measures the frequency with which newspapers use language that would tend to sway readers to the right or to the left on political issues. We focus on newspapers' news (rather than opinion) content, because of its centrality to public policy debates and its importance as a source of information to consumers.[2] To measure news slant, we examine the set of all phrases used by members of Congress in the 2005 *Congressional Record*, and identify those that are used much more frequently by one party than by another. We then index newspapers by the extent to which the use of politically charged phrases in their news coverage resembles the use of the same phrases in the speech of a congressional Democrat or Republican. The resulting index allows us to

[2]Nearly two-thirds of Americans report getting news several times a week or daily from local newspapers (Harris Interactive (2006)). Independent evidence suggests that almost 90 percent of readers of daily newspapers read the main news section, with over 80 percent reading the local news section (Newspaper Association of America (2006)).

compare newspapers to one another, though not to a benchmark of "true" or "unbiased" reporting.

Two key pieces of evidence suggest that our methodology produces a meaningful measure of slant. First, many of the phrases that our automated procedure identifies are known from other sources to be chosen strategically by politicians for their persuasive impact. Examples include "death tax," "tax relief," "personal account," and "war on terror" (which we identify as strongly Republican), and "estate tax," "tax break," "private account," and "war in Iraq," (which we identify as strongly Democratic). Second, the index that we construct using counts of these phrases in news coverage is consistent with readers' subjective evaluation of newspapers' political leanings (data on which are available for several large papers in our sample).

We use our measure to estimate a model of newspaper demand, in which a consumer's utility from reading a newspaper depends on the match between the newspaper's slant and the consumer's own ideology (Mullainathan and Shleifer (2005), Gentzkow and Shapiro (2006)). Using zip code-level data on newspaper circulation, we show that right-wing newspapers circulate relatively more in zip codes with a higher proportion of Republicans, even within a narrowly defined geographic market. Left-wing newspapers show the opposite pattern. Because we only use within-market variation to identify our model, our estimates are consistent even though slant is endogenous to the average political tastes in a market. We show that our results are also robust to correcting for measurement error (and for a subtler form of endogeneity bias) using an identification strategy in the spirit of George and Waldfogel (2003).

Treating newspapers as local monopolists, we compute the slant that each newspaper would choose if it independently maximized its own profits. The average profit-maximizing slant is close to the newspapers' actual slant. This finding is relevant to theories in which supply-side forces cause distortions in slant at the aggregate level. For example, if either the party identity of national incumbent politicians (Besley and Prat (2006)) or the distribution of political views among journalists in the country as a whole (Baron (2006)) were important drivers of slant, we would have expected to see deviation from profit maximization on average.

We also estimate a model of the supply of slant, in which we allow slant to respond both to the ideology of a newspaper's customers and also to the identity of its owner.

Variation in slant across newspapers is strongly related to the political makeup of their potential readers and thus to our estimated profit-maximizing points. The relationship between slant and consumer ideology remains when we compare different newspapers with the same owner or different newspapers in the same state. Overall, variation in consumer political attitudes explains roughly 20 percent of the variation in measured slant in our sample.

An obvious concern in interpreting the relationship between slant and consumer attitudes is that it may reflect causation running from slant to consumer

beliefs rather than the reverse. To address this, we show that the relationship survives when we instrument for consumer political attitudes using religiosity—a strong predictor of political preferences that is unlikely to be affected by newspaper content. These results do not mean that newspapers do not affect beliefs; indeed, our study is motivated in part by evidence that they do. Rather, our findings suggest that the effect of slant on ideology accounts for only a small part of the cross-sectional variation in ideology that identifies our model.

We find little evidence that the identity of a newspaper's owner affects its slant. After controlling for geographic clustering of newspaper ownership groups, the slant of co-owned papers is only weakly (and statistically insignificantly) related to a newspaper's political alignment. Direct proxies for owner ideology, such as patterns of corporate or executive donations to political parties, are also unrelated to slant. Estimates from a random effects model confirm a statistically insignificant role for owners, corresponding to approximately 4 percent of the variance in measured slant.

In the final section of the paper, we present additional evidence on the role of pressure from incumbent politicians (Besley and Prat (2006)), and the tastes of reporters and editors (Baron (2006)). The evidence we present suggests that neither of these forces is likely to explain a large share of the variation in slant.

This paper presents the first large-scale empirical evidence on the determinants of political slant in the news,[3] and informs the theoretical literature on demand-side (Mullainathan and Shleifer (2005), Gentzkow and Shapiro (2006), Suen (2004)) and supply-side (Besley and Prat (2006), Balan, DeGraba, and Wickelgren (2009), Baron (2006)) drivers of slant. Our findings contribute to the literature on product positioning in the mass media (Sweeting (2007, 2008), Myers (2008), George (2007)), as well as to research on product differentiation more generally (Mazzeo (2002a, 2002b), Dranove, Gron, and Mazzeo (2003), Seim (2006), Dubé, Hitsch, and Manchanda (2005), Einav (2007)).

Our work also advances the measurement of media slant (Groseclose and Milyo (2005), Puglisi (2008), Larcinese, Puglisi and Snyder (2007), Gentzkow, Glaeser, and Goldin (2006)).[4] Groseclose and Milyo (2005) use Congressional citations to estimate the political positions of think tanks, and then use data on media mentions of the same set of think tanks to measure the bias of 20 news outlets. Our automated procedure allows us to measure the slant of a much wider range of outlets, including over 400 daily newspapers representing over

[3]Hamilton (2004) presented an important overview of many of the issues we explore. An existing literature explores the determinants of newspaper endorsements of political candidates, rather than news content (see, e.g., Akhavan-Majid, Rife, and Gopinath (1991) or Ansolabehere, Lessem, and Snyder (2006)).

[4]Our approach borrows tools from the computer science literature on text categorization (see Aas and Eikvil (1999) for a review), which social scientists have applied to the measurement of sentiment (e.g., Antweiler and Frank (2004)) and politicians' platforms (Laver, Benoit, and Garry (2003)), but not (to our knowledge) to the political slant of the news media.

70 percent of total daily circulation in the United States. Moreover, rather than imposing a list of likely partisan phrases (such as names of think tanks), we use data from Congress to isolate the phrases that have the most power to identify the speaker's ideology.

The remainder of the paper is organized as follows. Section 2 discusses our data sources. Section 3 describes the computation of our measure of newspaper slant and validates the measure using alternative rankings of newspapers' political content. Section 4 presents our model, and Section 5 discusses identification and estimation. Sections 6, 7, and 8 present our core results. Section 9 tests two prominent theories of the determinants of media slant. Section 10 concludes.

2. DATA

2.1. *Congressional Record and Congressperson Data*

Our approach to measuring slant requires data on the frequency with which individual members of Congress use particular phrases. We use the text of the 2005 *Congressional Record*, downloaded from thomas.loc.gov and parsed using an automated script that identifies the speaker of each passage. To increase the efficiency of our text analysis algorithm, we apply a standard preprocessing procedure that removes extremely common words (such as "to," "from," and "the") and strips words down to shared linguistic roots (so that, for example, "tax cut" and "tax cuts" are identified as the same phrase). A final script produces counts by speaker and party of two- and three-word phrases in the *Congressional Record*. Appendix A contains additional details on this process.

For each congressperson (member of the House or Senate), we obtain data on party identification, as well as the share of the 2004 two-party presidential vote total going to George W. Bush in the congressperson's constituency (congressional district for representatives; state for senators). This vote share (which comes from polidata.org in the case of congressional districts) serves as our primary measure of a congressperson's ideology. We show in the online Appendix B (Gentzkow and Shapiro (2010)) that it is highly correlated with two commonly used roll-call measures of congressional ideology and that our results are robust to using these alternative measures of ideology as the basis for our analysis.

2.2. *Newspaper Text and Characteristics*

As an input to our slant measure, we obtain counts of the frequency with which phrases appear in news coverage from two sources: the NewsLibrary data base (newslibrary.com) and the ProQuest Newsstand data base (proquest.com). For each data base, we use an automated script to calculate the number of articles containing each phrase in each newspaper during calendar year 2005. Whenever possible, we exclude opinion content. Also, because

some newspapers do not archive reprinted wire stories with ProQuest, we exclude articles from the Associated Press, focusing instead on content originating with the newspaper. Appendix A provides additional details on the mechanics of these searches.

We compute slant for all English language daily newspapers available in either ProQuest or NewsLibrary for a total sample of 433 newspapers.[5] These newspapers together represented 74 percent of the total circulation of daily newspapers in the United States in 2001.

To measure the ownership and market characteristics of the newspapers in our sample, we first match every newspaper to data from the 2001 Editor and Publisher (E&P) International Yearbook CD-ROM. The E&P data set identifies the owner of each newspaper as of 2000.

The E&P data set also identifies the zip code of each newspaper's headquarters, which we match to counties using the United States 5-Digit ZIP Code Database from Quentin Sager Consulting. We match counties to primary metropolitan statistical areas (PMSAs) using definitions from the 1990 census. We define each newspaper's geographic market as the PMSA in which it is headquartered. If a newspaper is not located inside a PMSA, we define its market to be the county in which it is located. For the median newspaper, this market definition includes more than 90 percent of the newspaper's total circulation (among newspapers for which we have zip code-level circulation data). For four newspapers—the *New York Times*, the *Wall Street Journal*, the *Christian Science Monitor*, and *USA Today*—the notion of a geographic market is ill defined. We exclude these papers from our analysis, leaving a sample of 429 newspapers with well defined geographic markets.

For each newspaper, we obtain a wide range of demographic characteristics of the paper's market from the 2000 U.S. Census. We also obtain data from David Leip's Atlas of U.S. Presidential Elections (uselectionatlas.org) on the share of votes in each market going to Bush in the 2004 presidential election; this is used as a proxy for the market's political leanings. Last, we use the DDB Needham Life Style Survey (Putnam (2000)), available on bowlingalone.com, to compute a measure of the share of survey respondents from 1972 to 1998 who reported attending church monthly or more. This measure serves as a plausibly exogenous shifter of the political leanings of the market in that it is unlikely to be directly affected by the slant of area newspapers.

As a potential proxy for a media firm's ideological leanings, we obtain data from the Center for Public Integrity (publicintegrity.org) on the share of each newspaper firm's corporate political contribution dollars going to Republicans. We also searched the Federal Election Commission (FEC) disclosure data base

[5]One additional newspaper—the *Chicago Defender*—is present in the news data bases, but is excluded from our analysis because it is an extreme outlier (more than 13 standard deviations away from the mean) in the distribution of slant. A large share of hits for this paper are for a single phrase, "African American," which is strongly predictive of liberal ideology in Congress.

for information on the personal contributions of the Chief Executive Officer, President, Chairman, and Managing Director of each firm that owns two or more U.S. daily newspapers. For newspapers owned by a firm with no other daily newspaper holdings, we conducted an analogous search, but collected data on executives of the newspaper itself.

2.3. *Newspaper Circulation and Consumer Characteristics*

For our study of the effects of slant on newspaper demand, we use zip code-level data on newspaper circulation from the Audit Bureau of Circulation's (ABC) Newspaper GeoCirc data set. We include all zip code–newspaper pairs with positive circulation. We match each zip code to a news market using the market definition above.

To adjust for nonpolitical differences across zip codes, we make use of a set of zip code demographics taken from the 2000 U.S. Census (census.gov): log of total population, log of income per capita, percent of population urban, percent white, percent black, population per square mile, share of houses that are owner occupied, and the share of population 25 and over whose highest level of schooling is college.

Measuring each zip code's ideology is complicated by the fact that voting data are not available at the zip code level. To circumvent this problem, we use the Federal Election Commission's (FEC) 2000, 2002, and 2004 Individual Contributions Files. These files, which are available for download at fec.gov, contain a record of every individual contribution to a political party, candidate, or political action committee registered with the FEC. Each donor record includes a complete address, allowing us to identify donors' zip codes. For each zip code, we compute the share of donations (denominated in number of donations, not dollars) received by a Republican affiliate among donations received by either Republican- or Democrat-affiliated entities. To reduce the noise in the measure, we restrict attention to zip codes with 20 or more donors.

To test the validity of this proxy for ideology, we take advantage of data on the number of registered Democrats and Republicans by zip code in California as of March 2006.[6] The donation measure has a correlation of 0.65 with the two-party share of Republican registrants.

Of course, the sample of donors to political causes is not fully representative of the entire population of a zip code. Donors tend to be older, richer, and more educated than nondonors (Gimpel, Lee, and Kaminski (2006)). However, these are also the demographic characteristics of likely readers of newspapers (Gentzkow (2007)) and, therefore, if anything, may tend to make our measure more representative of the population relevant for studying newspaper demand.

[6]We are grateful to Marc Meredith for providing these data.

Our analysis of newspaper demand is restricted to the 290 newspapers in our primary sample for which we observe at least one zip code with both positive circulation in the ABC data and sufficiently many donors in the FEC data.

3. MEASURING SLANT

Our approach to measuring the slant of a newspaper will be to compare phrase frequencies in the newspaper with phrase frequencies in the 2005 *Congressional Record* to identify whether the newspaper's language is more similar to that of a congressional Republican or a congressional Democrat.

For a concrete illustration of our approach to measuring slant, consider the use of the phrases "death tax" and "estate tax" to describe the federal tax on assets of the deceased. The phrase "death tax" was coined by the tax's conservative opponents. According to a high-level Republican staffer, "Republicans put a high level of importance on the death/estate tax language—they had to work hard to get members to act in unison, including training members to say 'death tax'... Estate tax sounds like it only hits the wealthy but 'death tax' sounds like it hits everyone" (Graetz and Shapiro (2005)). In Congress in 2005, Republicans used the phrase "death tax" 365 times and the phrase "estate tax" only 46 times. Democrats, by contrast, had the reverse pattern, using the phrase "death tax" only 35 times and the phrase "estate tax" 195 times.

The relative use of the two phrases in newspaper text conforms well to prior expectations about political slant. Compare, for example, the *Washington Post* and the *Washington Times.* The *Post* is widely perceived to be more liberal than the *Times*.[7] In 2005, the *Post* used the phrase "estate tax" 13.7 times as often as it used the phrase "death tax," while the *Times* used "estate tax" 1.3 times as often. As we show below, this case is not unusual: there is a significant correlation between popular perceptions of a newspaper's political leanings and its propensity to use words and phrases favored by different political parties in Congress. Our measure of media slant exploits this fact by endogenously identifying politically charged phrases like "death tax" and "estate tax," and computing their frequencies in daily newspapers throughout the United States.

In principle, we could base our measure on counts of *all* phrases that appear in the *Congressional Record*. A simple procedure would be as follows. First, for each politician, we compute a vector that gives the number of times each phrase appeared in their speeches. Second, we compute a mapping from the vector of counts to a measure of a politician's ideology. Finally, we generate counts of each phrase in a newspaper's text and apply the same mapping to generate an index of the newspaper's ideology.

[7]The website mondotimes.com presents an index of newspapers' political leanings based on user ratings. The *Post* is rated as "leans left," while the *Times* is rated as "conservative." Groseclose and Milyo (2005) also rated the *Post* as significantly to the left of the *Times*.

Because the total number of phrases that appear in the *Congressional Record* is in the millions, this simple procedure is computationally infeasible. We therefore add a "feature selection" step in which we use simple computations to identify a set of phrases that are highly diagnostic of the speaker's political party. We use this restricted phrase set for the more computationally burdensome step of mapping phrase counts to a continuous measure of ideology, counting occurrences in newspapers, and estimating newspaper ideology.

3.1. *Selecting Phrases for Analysis*

Let f_{pld} and f_{plr} denote the total number of times phrase p of length l (two or three words) is used by Democrats and Republicans, respectively. Let $f_{\sim pld}$ and $f_{\sim plr}$ denote the total occurrences of length-l phrases that are *not* phrase p spoken by Democrats and Republicans, respectively. Let χ^2_{pl} denote Pearson's χ^2 statistic for each phrase:

$$\chi^2_{pl} = \frac{(f_{plr}f_{\sim pld} - f_{pld}f_{\sim plr})^2}{(f_{plr} + f_{pld})(f_{plr} + f_{\sim plr})(f_{pld} + f_{\sim pld})(f_{\sim plr} + f_{\sim pld})}. \tag{1}$$

We select the phrases for our analysis as follows:

(i) We compute the total number of times that each phrase appeared in newspaper headlines and article text in the ProQuest Newsstand data base from 2000 to 2005. We restrict attention to two-word phrases that appeared in at least 200 but no more than 15,000 newspaper headlines, and three-word phrases that appeared in at least 5 but no more than 1000 headlines. We also drop any phrase that appeared in the full text of more than 400,000 documents.

(ii) Among the remaining phrases, we select the 500 phrases of each length l with the greatest values of χ^2_{pl}, for a total of 1000 phrases.

The first step eliminates phrases that are not likely to be useful for diagnosing newspaper partisanship. For example, procedural phrases such as "yield the remainder of my time" are commonly employed in the *Congressional Record*—especially by the majority party—but are almost never used in newspapers. Extremely common phrases such as "third quarter" or "exchange rate" are also unlikely to be diagnostic of ideology, but impose a high burden on our procedure for extracting phrase counts in newspaper text. The cutoffs we impose are arbitrary. In (online) Appendix B, we show that our results are robust to tightening these cutoffs.

The second step identifies phrases that are diagnostic of the speaker's political party. If the counts f_{pld} and f_{plr} are drawn from (possibly different) multinomial distributions, χ^2_{pl} is a test statistic for the null hypothesis that the propensity to use phrase p of length l is equal for Democrats and Republicans. This statistic conveniently summarizes the political asymmetry in the use of the phrase. (More naive statistics, such as the ratio of uses by Republicans to uses by Democrats, would tend to select phrases that are used only once by

TABLE I

MOST PARTISAN PHRASES FROM THE 2005 *CONGRESSIONAL RECORD*[a]

Panel A: Phrases Used More Often by Democrats		
Two-Word Phrases		
private accounts	Rosa Parks	workers rights
trade agreement	President budget	poor people
American people	Republican party	Republican leader
tax breaks	change the rules	Arctic refuge
trade deficit	minimum wage	cut funding
oil companies	budget deficit	American workers
credit card	Republican senators	living in poverty
nuclear option	privatization plan	Senate Republicans
war in Iraq	wildlife refuge	fuel efficiency
middle class	card companies	national wildlife
Three-Word Phrases		
veterans health care	corporation for public broadcasting	cut health care
congressional black caucus	additional tax cuts	civil rights movement
VA health care	pay for tax cuts	cuts to child support
billion in tax cuts	tax cuts for people	drilling in the Arctic National
credit card companies	oil and gas companies	victims of gun violence
security trust fund	prescription drug bill	solvency of social security
social security trust	caliber sniper rifles	Voting Rights Act
privatize social security	increase in the minimum wage	war in Iraq and Afghanistan
American free trade	system of checks and balances	civil rights protections
central American free	middle class families	credit card debt

(*Continues*)

Republicans and never by Democrats, even though pure sampling error could easily generate such a pattern.) χ^2_{pl} is also simple to compute, in the sense that it requires only two calculations per phrase: the number of uses by Republicans and the number of uses by Democrats.

Table I shows the top phrases (arranged in order of descending χ^2_{pl} by length) in our final set of 1000. Panel A shows phrases used more often by congressional Democrats. Panel B shows phrases used more often by congressional Republicans.

Our procedure identifies many phrases that both intuition and existing evidence suggest are chosen strategically for their partisan impact. For example, a widely circulated 2005 memo by Republican consultant Frank Luntz advised candidates on the language they should use to describe President Bush's proposed Social Security reform (Luntz (2005)):

> Never say 'privatization/private accounts.' Instead say 'personalization/personal accounts.' Two-thirds of America want to personalize Social Security while only one-third would privatize it. Why? Personalizing Social Security suggests ownership and control over your retirement savings, while privatizing it suggests a profit motive and winners and losers.

TABLE I—*Continued*

Panel B: Phrases Used More Often by Republicans		
Two-Word Phrases		
stem cell	personal accounts	retirement accounts
natural gas	Saddam Hussein	government spending
death tax	pass the bill	national forest
illegal aliens	private property	minority leader
class action	border security	urge support
war on terror	President announces	cell lines
embryonic stem	human life	cord blood
tax relief	Chief Justice	action lawsuits
illegal immigration	human embryos	economic growth
date the time	increase taxes	food program
Three-Word Phrases		
embryonic stem cell	Circuit Court of Appeals	Tongass national forest
hate crimes legislation	death tax repeal	pluripotent stem cells
adult stem cells	housing and urban affairs	Supreme Court of Texas
oil for food program	million jobs created	Justice Priscilla Owen
personal retirement accounts	national flood insurance	Justice Janice Rogers
energy and natural resources	oil for food scandal	American Bar Association
global war on terror	private property rights	growth and job creation
hate crimes law	temporary worker program	natural gas natural
change hearts and minds	class action reform	Grand Ole Opry
global war on terrorism	Chief Justice Rehnquist	reform social security

[a]The top 60 Democratic and Republican phrases, respectively, are shown ranked by χ^2_{pl}. The phrases are classified as two or three word after dropping common "stopwords" such as "for" and "the." See Section 3 for details and see Appendix B (online) for a more extensive phrase list.

We identify "personal accounts," "personal retirement accounts," and "personal savings accounts" as among the most Republican phrases in the *Congressional Record*, while "private accounts," "privatization plan," and other variants show up among the most Democratic phrases. Similarly, we identify "death tax" (whose partisan pedigree we discussed above) as the third most Republican phrase. We identify "tax relief"—a term also advocated by Luntz (2005)—as strongly Republican, while "tax breaks" is strongly Democratic. On foreign policy, we identify variants on the phrase "global war on terror" as among the most strongly Republican phrases, while "war in Iraq" and "Iraq war" are Democratic, again consistent with accounts of party strategy (e.g., Stevenson (2005)).

The phrases in our sample arise regularly in news content. The average newspaper in our sample used these phrases over 13,000 times in 2005. Even newspapers in the bottom quartile of daily circulation (in our newspaper sample) use these phrases over 4000 times on average. The contexts in which these phrases appear include local analogues of national issues, local impact of federal legislation, and the actions of legislators from local districts. In Ap-

pendix A, we present more systematic evidence on the contexts in which our phrases appear. Most occurrences are in independently produced news stories.

3.2. *Mapping Phrases to Ideology*

Re-index the phrases in our sample by $p \in \{1, \ldots, 1000\}$. (Ignore phrase length for notational convenience.) For each congressperson $c \in C$, we observe ideology y_c and phrase frequencies $\{f_{pc}\}_{p=1}^{1000}$. Let $\tilde{f}_{pc} \equiv f_{pc} / \sum_{p=1}^{P} f_{pc}$ denote the relative frequency of phrase p in the speech of congressperson c.

We have a set of newspapers $n \in N$ for which we observe phrase frequencies $\{f_{pn}\}_{p=1}^{1000}$ but not ideology y_n. We estimate ideology for newspapers as follows:

(i) For each phrase p, we regress $\tilde{f}_{pc}$ on y_c for the sample of congresspeople, obtaining intercept and slope parameters a_p and b_p, respectively.

(ii) For each newspaper n, we regress $(\tilde{f}_{pn} - a_p)$ on b_p for the sample of phrases, obtaining slope estimate

$$\hat{y}_n = \frac{\sum_{p=1}^{1000} b_p (\tilde{f}_{pn} - a_p)}{\sum_{p=1}^{1000} b_p^2}. \tag{2}$$

(We also compute an analogous estimate $\hat{y}_c$ for each congressperson c.)

This approach can be understood as follows. First, we use congresspeople—whose ideology is observed—to estimate the relationship between the use of a phrase p and the ideology of the speaker. Second, we use the relationship observed in the first stage to infer the ideology of newspapers by asking whether a given newspaper tends to use phrases favored by more Republican members of Congress. If the use of some phrase p is uncorrelated with a congressperson's ideology ($b_p = 0$), the use of that phrase does not contribute to the estimate $\hat{y}_n$. If phrase p is used more often by more right-wing congresspeople ($b_p > 0$), the estimator will judge a speaker who uses phrase p often as more right wing. If newspaper phrase frequencies are given by $\tilde{f}_{pn} = a_p + b_p y_n + e_{pn}$, with $E(e_{pn} \mid b_p) \equiv 0 \; \forall n$, then $E(\hat{y}_n) = y_n \; \forall n$.

The estimates $\hat{y}_c$ have a correlation of 0.61 with true ideology y_c among our sample of congresspeople. This correlation provides in-sample evidence for the validity of our estimates, but also implies that our estimates are likely to contain a significant amount of noise. Taking the square of the correlation coefficient, 37 percent of the variation in slant is attributable to variation in ideology, with the rest coming from noise. Therefore, a useful benchmark is that, assuming the same share of noise among congresspeople and newspapers, 63 percent of the variation in slant among newspapers is likely to be noise.

Validating our approach among newspapers is more difficult. The estimate $\hat{y}_n$ attempts to answer the question, "If a given newspaper were a congressperson,

how Republican would that congressperson's district be?" By definition, the true answer to this question is unobservable for newspapers, but a crude proxy is available. The media directory website Mondo Times (mondotimes.com) collects ratings of newspapers' political orientation from its users.[8] Note that we would not necessarily expect these correlations to be perfect, both because most papers receive only a few ratings and because Mondo Times users are rating the opinion as well as news content of the papers, whereas our slant measure focuses on news content. Nevertheless, in Figure 1 we show that these

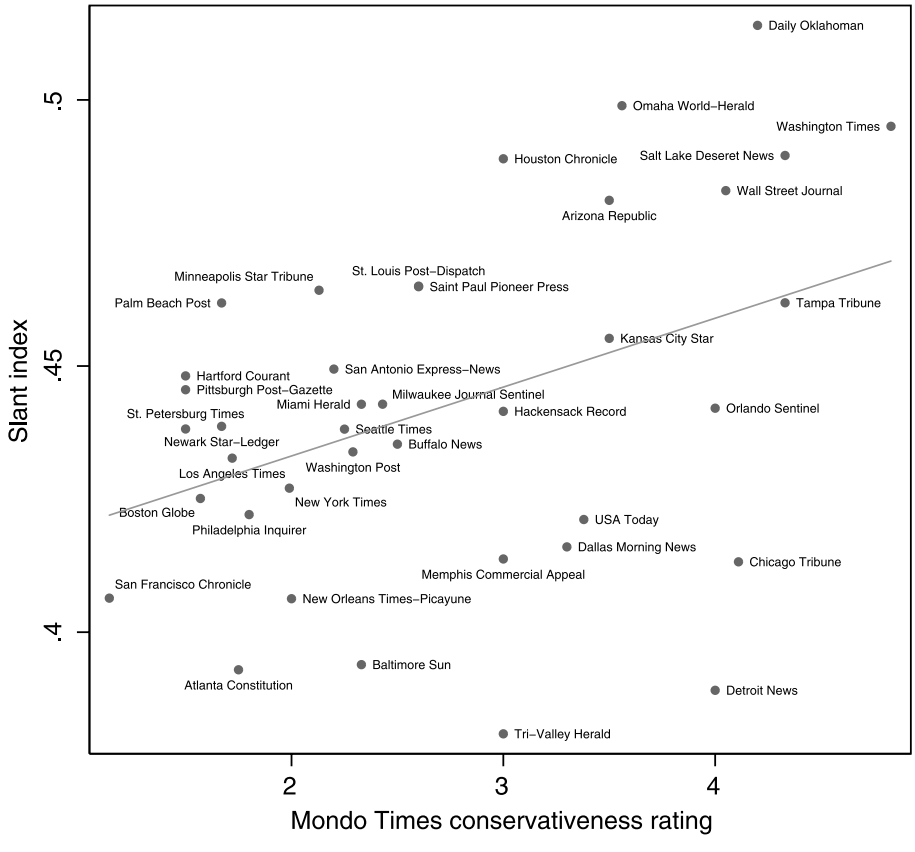

FIGURE 1.—Language-based and reader-submitted ratings of slant. The slant index (y axis) is shown against the average Mondo Times user rating of newspaper conservativeness (x axis), which ranges from 1 (liberal) to 5 (conservative). Included are all papers rated by at least two users on Mondo Times, with at least 25,000 mentions of our 1000 phrases in 2005. The line is predicted slant from an OLS regression of slant on Mondo Times rating. The correlation coefficient is 0.40 ($p = 0.0114$).

[8]We wish to thank Eric Kallgren of Mondo Code for graciously providing these data.

ratings are positively related to our slant index with a correlation coefficient of 0.40.

4. AN ECONOMIC MODEL OF SLANT

In this section we define the demand for and supply of slant. Our model is designed to capture three important features of newspaper markets. First, consumers may prefer newspapers whose slant is close to their own ideology. Second, firms will have an incentive to cater to this demand. Finally, owner ideology may also affect firms' choices of slant and this may lead slant to differ from the profit-maximizing level.

4.1. *Consumer Problem*

Each zip code z contains a continuum of households of mass H_z, with individual households indexed by i. A set of newspapers N_z is available in each zip code, and each household i must choose a subset $N_{iz} \subseteq N_z$ of the available newspapers to read. Household i in zip code z gets value u_{izn} from reading newspaper n, and the utility U_{iz} of household i is

$$U_{iz} \equiv \sum_{n \in N_{iz}} u_{izn}.$$

Consistent with utility maximization, household i in zip code z reads newspaper $n \in N_z$ iff $u_{izn} \geq 0$.

Each zip code z has an exogenous ideology r_z (with higher values meaning more conservative) and a preferred slant

$$\text{ideal}_z \equiv \alpha + \beta r_z.$$

If $\beta > 0$, more conservative zip codes prefer more conservative news, as in Mullainathan and Shleifer (2005).

Household utility u_{izn} is the sum of three components:

$$u_{izn} \equiv \bar{u}_{zn} - \gamma(y_n - \text{ideal}_z)^2 + \varepsilon_{izn}. \tag{3}$$

The term $\bar{u}_{zn}$ is the exogenous taste of consumers in zip code z for newspaper n, possibly related to observables, but not affected by slant y_n. The term $-\gamma(y_n - \text{ideal}_z)^2$ captures the distaste for reading a newspaper whose slant y_n deviates from the preferred slant ideal_z. The error term ε_{izn} is a household-specific taste shock which we assume has a logistic distribution. We assume that $\bar{u}_{zn}$ is known to firms (but not necessarily to the econometrician).

The share of households in zip code z reading newspaper n is then

$$S_{zn} = \frac{\exp[\bar{u}_{zn} - \gamma(y_n - \text{ideal}_z)^2]}{1 + \exp[\bar{u}_{zn} - \gamma(y_n - \text{ideal}_z)^2]} \tag{4}$$

if $n \in N_z$ and 0 otherwise.

If $\gamma, \beta > 0$, it is straightforward to show that equation (4) implies two key testable hypotheses:

HYPOTHESIS D1: Fixing $\bar{u}_{zn}$

$$\frac{\partial^2}{\partial y_n \partial r_z}\left(\ln \frac{S_{zn}}{1 - S_{zn}}\right) > 0.$$

More conservative zip codes have a relatively greater taste for more conservatively slanted news.

HYPOTHESIS D2: Fixing $\bar{u}_{zn}$ and y_n

$$\frac{\partial^2}{\partial r_z^2}\left(\ln \frac{S_{zn}}{1 - S_{zn}}\right) < 0.$$

Demand has an inverted-U relationship to zip code ideology, peaking at $r_z = (y_n - \alpha)/\beta$.

4.2. *Firm Problem*

Assume that circulation revenue, advertising revenue, and variable costs are proportional to circulation, so that each newspaper earns a fixed markup for each copy sold. Let ideal_n be the value of y_n that maximizes newspaper n's circulation. If all newspapers were operated by profit-maximizing firms, equilibrium slant would be $y_n^* = \text{ideal}_n$.

We allow for deviations from profit maximization. Each newspaper n is owned by a firm g, which has an ideology μ_g. Equilibrium slant is given by

$$y_n^* = \rho_0 + \rho_1 \text{ideal}_n + \mu_g. \tag{5}$$

When $\rho_0 = 0$, $\rho_1 = 1$, and $\mu_g = 0$, equation (5) is equivalent to profit maximization. Equation (5) can therefore be thought of as an approximation to a model in which a newspaper owner maximizes a utility function that includes dollar profits as well as nonpecuniary ideological motivations. In Gentzkow and Shapiro (2007), we derived an expression analogous to equation (5) from a set of primitive assumptions on consumers' and firms' utility functions.

We highlight two testable hypotheses of the model:

HYPOTHESIS S1: $\partial y_n / \partial \text{ideal}_n > 0$. Slant is increasing in consumer Republicanism.

HYPOTHESIS S2: $\partial y_n / \partial \mu_g > 0$. Slant is increasing in owner Republicanism.

4.3. *Discussion*

Our model is restrictive in a number of respects.

First, we do not explicitly model the fact that consumer ideology r_z may itself be a function of slant. Evidence suggests that slant does affect political behavior; this is an important motivation for our study. However, we expect that *most* of the variation in consumer ideology is related to consumer characteristics such as geography, race, and religiosity that are not affected by newspapers, making the potential bias in our estimates from ignoring reverse causality relatively small. In Section 7.1, we support this interpretation directly using an instrumental variables strategy in a cross-market regression of slant on consumer ideology. It is worth stressing, however, that we do not have an analogous instrument for the within-market (cross zip code) variation in ideology that identifies our demand model. Our demand estimates therefore rely more heavily than our supply estimates on the assumption that most variation in ideology is exogenous with respect to newspaper content.

Second, we assume that ideology does not vary across consumers within a zip code. This assumption approximates a model in which the average Republican in a heavily Republican zip code is further to the right than the average Republican in a more liberal zip code. In Gentzkow and Shapiro (2007), we showed that our main findings survive in a model that allows explicitly for within-zip code heterogeneity in political ideology.

Third, we assume that consumer utility is additive over newspapers, thus eliminating complementarity or substitutability in demand, and ruling out strategic interactions among newspapers. Since only a handful of papers in our sample face same-city competitors, we view a model without strategic interactions as a reasonable approximation. Excluding newspapers with same-city competitors does not change our results regarding the supply of slant (see online Appendix B). Our model does, however, ignore some potentially important strategic interactions, such as between newspapers and local television stations or newspapers in neighboring cities.

Fourth, we normalize the outside option to zero for all consumers. The outside option captures the value of all alternatives not written into the model, including television news, Internet news, and so forth. Because we will include market–newspaper fixed effects (FE) in our demand estimation, we in fact allow the utility of the outside option to vary nonparametrically by market. We do not, however, allow its utility to vary across zip codes; in particular, we rule out variation that is correlated with r_z. That assumption is important for our tests of Hypothesis D2 and for our structural estimates. It is not important for our tests of Hypothesis D1 and, indeed, we find evidence for Hypothesis D1 in a zip code fixed effects specification that allows arbitrary variation in the outside option across zip codes.

Finally, we assume that the markup newspapers earn is the same for each unit of circulation, whereas in reality advertisers prize some readers more than others. We show in the online Appendix B that allowing advertising revenues

per reader to vary across zip codes as a function of demographic characteristics does not change our conclusions.

5. IDENTIFICATION AND ESTIMATION

5.1. *Demand Parameters*

To estimate the demand model of equation (3), we specify the zip code–newspaper taste parameter $\bar{u}_{zn}$ as

$$\bar{u}_{zn} = X_z\phi^0 + W_{zn}\phi^1 + \xi_{mn} + \nu_{zn}, \tag{6}$$

where ϕ^0 and ϕ^1 are parameter vectors, X_z is a vector of zip code demographics, W_{zn} is a vector of interactions between the zip code demographics in X_z and the average level of the corresponding demographics in the newspaper's market, ξ_{mn} is an unobservable product characteristic that is allowed to vary at the market level, and ν_{zn} is a zip code–newspaper-level unobservable.

Substituting for $\bar{u}_{zn}$ and ideal$_z$ in equation (4), and combining terms that do not vary within market–newspaper pairs, we have our estimating equation

$$\ln \frac{S_{zn}}{1 - S_{zn}} = \delta_{mn} + \lambda_0^d y_n r_z + \lambda_1^d r_z + \lambda_2^d r_z^2 + X_z\phi^0 + W_{zn}\phi^1 + \nu_{zn}, \tag{7}$$

where $\lambda_0^d = 2\gamma\beta$, $\lambda_1^d = -2\gamma\alpha\beta$, and $\lambda_2^d = -\gamma\beta^2$, and where we treat the market–newspaper term

$$\delta_{mn} = -\gamma\alpha^2 - \gamma y_n^2 + 2\gamma\alpha y_n + \xi_{mn}, \tag{8}$$

as a fixed effect.

We adopt an instrumental variables strategy to allow for measurement error in $\hat{y}_n$. We let R_n be the overall share of Republicans in newspaper n's primary market, measured using the Republican share of the 2004 two-party vote for president. We make the following assumptions:

(i) $E[(\hat{y}_n - y_n) \mid R_n, r_z, X_z, W_{zn}, \delta_{mn}] = 0$.
(ii) $E[\nu_{zn} \mid R_n, r_z, X_z, W_{zn}, \delta_{mn}] = 0$.

Under these assumptions, we consistently estimate the parameters of equation (7) via two-stage least squares, treating $r_z\hat{y}_n$ as an endogenous regressor, r_zR_n as an excluded instrument, and δ_{mn} as a fixed effect. We allow for correlation in the error term ν_{zn} across observations for a given newspaper n.

Our instrumental variables strategy builds on George and Waldfogel's (2003) insight that because fixed costs lead newspapers to cater to the average tastes of their readers, individuals will tend to read more when their tastes are similar to the average. By the same logic, our model predicts that if slant is an important component of demand, (i) newspapers with high R_n should choose high values of y_n and (ii) newspapers with high R_n should consequently be read relatively

more in zip codes with high r_z. The strength of these relationships will identify the coefficient on $y_n r_z$. Note that assuming that R_n is correlated with y_n is not equivalent to assuming that $y_n = y_n^*$ or that $y_n = \text{ideal}_n$. That is, for the purposes of our demand analysis, we do not assume that slant is chosen to maximize profits, only that it is correlated with consumer ideology in the newspaper's home market.

This strategy requires that the noise in our search-based measure of slant is unrelated to the characteristics of a newspaper's market. It also requires that we have controlled for zip code-specific factors that affect demand and are correlated with r_z or the interaction $r_z R_n$. Note that we do not need to assume that the market–newspaper taste shock ξ_{mn} is orthogonal to R_n: we allow for ξ_{mn} to be endogenous to R_n by treating δ_{mn} as a fixed effect.

Although our main reason for instrumenting is to correct for measurement error in $\hat{y}_n$, our instrument also addresses a subtle form of endogeneity bias. Note that the most obvious kind of endogeneity—that slant y_n may be a function of the unobserved product characteristic ξ_{mn}—would not affect even ordinary least squares (OLS) estimates because both the main effect of y_n and the unobservable ξ_{mn} are absorbed in δ_{mn}. However, slant could be endogenous, not to overall demand for the newspaper, but to the correlation between zip code ideology r_z and demand. More precisely, if the error term were written as $\tilde{\xi}_{mn} r_z + \nu_{zn}$, where $\tilde{\xi}_{mn}$ is a random coefficient, then slant y_n might tend to be higher for newspapers receiving a higher draw of $\tilde{\xi}_{mn}$, because such newspapers have (exogenously) greater presence in highly Republican zip codes. Such a force would bias OLS estimates upward (absent measurement error), but would be addressed by our instrumental variables strategy provided that $E[\tilde{\xi}_{mn} \mid R_n, r_z, X_z, W_{zn}, \delta_{mn}] = 0$.

Our controls address a range of other possible confounds. Including fixed effects δ_{mn} at the market–newspaper level will control for unobserved newspaper characteristics, unobserved market-level tastes, and heterogeneity in the "fit" between the newspaper and the market (say, because of physical distance). Zip code-level controls X_z account for the fact that demographics like education and race affect readership and may be correlated with political tastes. The interactions W_{zn} account for the fact that these other characteristics may have different effects on readership depending on the average characteristics of a newspaper's market (George and Waldfogel (2003)). For example, the percent black in a zip code may relate positively to readership of newspapers from predominantly black markets, and negatively on readership of newspapers from predominantly white neighborhoods.

5.2. *Supply Parameters*

To estimate the supply model of equation (5), we assume that true slant $y_n = y_n^*$, but allow that measured slant $\hat{y}_n \neq y_n$.

Because we can only calculate the profit-maximizing level of slant ideal_n directly for the 290 of newspapers in our demand sample, we approximate ideal_n as a linear function of the Republican vote share in a newspaper's market: $\widehat{\text{ideal}}_n = \eta_0 + \eta_1 R_n + \zeta_n$. This allows us to use our complete sample of 429 newspapers for the supply analysis.

Substituting $\widehat{\text{ideal}}_n$ in place of ideal_n, we then have the estimating equation

$$\hat{y}_n = \lambda_0^s + \lambda_1^s R_n + \mu_g + \omega_n, \tag{9}$$

where $\lambda_0^s = \rho_0 + \rho_1\eta_0$, $\lambda_1^s = \rho_1\eta_1$, and $\omega_n = \rho_1\zeta_n + (\hat{y}_n - y_n)$.

We assume that $\omega_n \sim N(\theta_s, \sigma_\omega^2)$, where s is the newspaper's home state. Here, θ_s is a state-specific measurement error component, with $E(\theta_s) \equiv 0$. We assume that $\mu_g \sim N(\bar{\mu}, \sigma_\mu^2)$, with μ_g, R_n, and ω_n orthogonal conditional on θ_s.

Equation (9) is then a random effects (RE) model. We will control for θ_s flexibly using state fixed effects. Variation in slant that is common to newspapers with the same owner is attributed to variation in μ_g. Newspaper-level variation that is not correlated across newspapers with the same owner is attributed to variation in ω_n.

We include the state-specific measurement error component θ_s in the model because the strong geographic clustering of ownership groups (Lacy and Simon (1997), Martin (2003)) means that any geographic component of measurement error, due to regional patterns of speech or news, could otherwise be spuriously attributed to owner tastes. Inclusion of this component means that variation in owner tastes is identified from correlation in deviations across newspapers with the same owner, after accounting for state effects. Identification therefore relies on the significant number of owners with geographically diverse holdings. Half of the ownership groups with multiple papers in our sample span more than two states. For example, the markets where the New York Times Company owns newspapers range from New York City to Sarasota, FL and Spartanburg, SC.

Our main specifications require that there is no causality running from $\hat{y}_n$ to R_n. We address the possibility of reverse causality below by instrumenting for R_n with consumer religiosity—a characteristic we expect to be a strong predictor of R_n but unaffected by $\hat{y}_n$.

6. EVIDENCE ON THE DEMAND FOR SLANT

Figure 2 presents evidence on Hypothesis D1. For each newspaper, we regress demand $\ln(S_{zn}/(1 - S_{zn}))$ on zip code ideology r_z, with fixed effects for market. We plot the resulting coefficients against measured slant $\hat{y}_n$ for the 59 newspapers that circulate in markets containing more than 200 zip codes (where coefficients are reasonably well identified). As predicted, the estimated

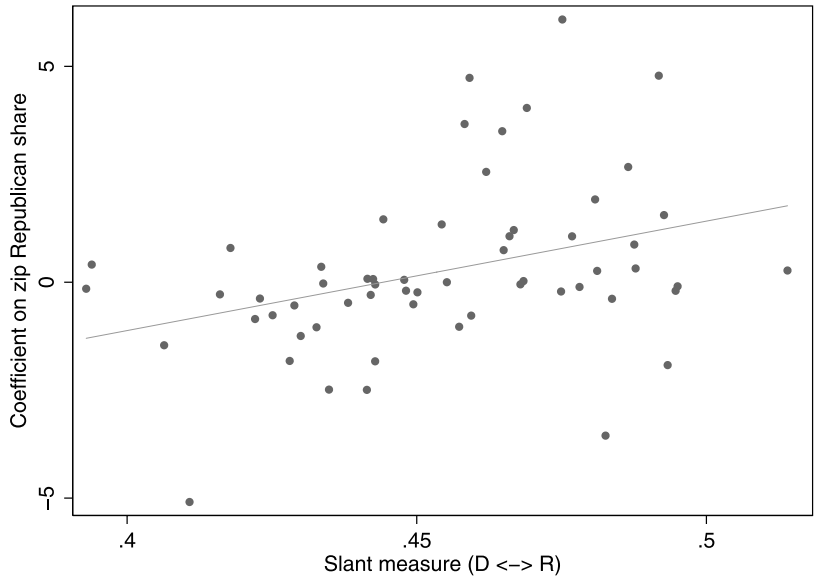

FIGURE 2.—Newspaper slant and coefficients on zip code ideology. The y axis shows the estimated coefficient in a regression of the share of households in the zip code reading each newspaper on the zip code share Republican, for newspapers circulating in more than 200 zip codes. The x axis shows slant measure.

effect of zip code Republicanism on demand has a clear positive relationship with the newspaper's slant.

Figure 3 presents evidence on Hypothesis D2. Each panel shows, for newspapers in a given quartile of the distribution of measured slant $\hat{y}_n$, the coefficients on dummies for deciles of zip code ideology r_z, in a regression of demand on decile dummies and market–newspaper fixed effects, weighted by H_z. The graphs are noisy but consistent with an inverted-U relationship, peaking further to the right at higher values of $\hat{y}_n$.

The first column of Table II presents these findings quantitatively. We regress $\ln(S_{zn}/(1 - S_{zn}))$ on $r_z\hat{y}_n$, r_z, and r_z^2, and adjust standard errors for correlation at the newspaper level. Consistent with Hypothesis D1, the coefficient on the interaction term $r_z\hat{y}_n$ is positive and statistically significant. Consistent with Hypothesis D2, the coefficient on r_z is negative and statistically significant, and the coefficient on r_z^2 is negative and marginally statistically significant.

The second column of Table II adds controls for zip code demographics X_z and zip code demographics interacted with market demographics W_{zn}. Our findings survive and, if anything, the evidence for Hypothesis D2 becomes stronger statistically.

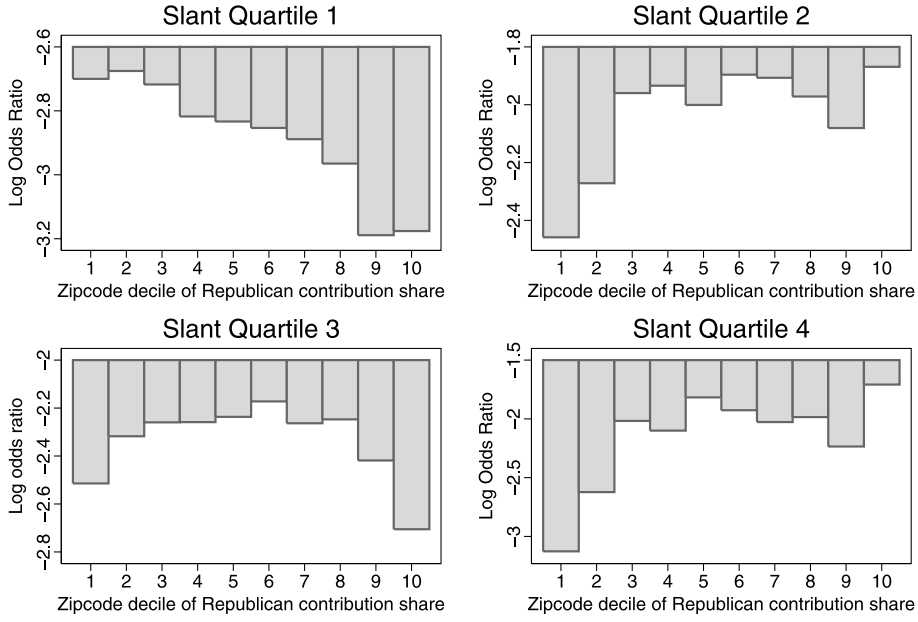

FIGURE 3.—Newspaper demand and zip code ideology by quartiles of newspaper slant. The coefficients on decile dummies in regressions of the share of households in a zip code reading a newspaper on dummies for decile of share donating to Republicans in the 2000–2004 election cycle are shown with market–newspaper fixed effects and weighted by zip code population. The equation is estimated separately for newspapers in each quartile of the distribution of measured slant.

The third column of Table II adds controls for zip code fixed effects. This model is identified from zip codes where two or more newspapers circulate. It allows for unobserved zip code characteristics that affect the overall propensity to read newspapers. In particular, it allows for the possibility that the utility of the outside option varies across zip codes in a way that is correlated with r_z. By definition, we cannot test Hypothesis D2 in this specification, but the evidence for Hypothesis D1 survives.

The last column of Table II presents estimates of our preferred demand model—estimating equation (7) under the assumptions of Section 5.1. We instrument for $r_z\hat{y}_n$ with $r_z R_n$ to address measurement error in $\hat{y}_n$. As expected, the coefficient on $r_z\hat{y}_n$ increases. The change in magnitude is quantitatively plausible: given that about 63 percent of the variation in $\hat{y}_n$ is measurement error, we would expect its coefficient to be attenuated by a factor of $\frac{1}{1-0.63} \approx 2.7$. In fact, the coefficient in the last column is about 2.6 times that in the second column.

TABLE II

EVIDENCE ON THE DEMAND FOR SLANT[a]

Description	Model			
	OLS	OLS	OLS	2SLS
(Zip share donating to Republicans) × Slant	10.66	9.441	14.61	24.66
	(3.155)	(2.756)	(6.009)	(7.692)
Zip share donating to Republicans	−4.376	−3.712	—	−10.41
	(1.529)	(1.274)		(3.448)
(Zip share donating to Republicans)^2	−0.4927	−0.5238	—	−0.7103
	(0.2574)	(0.2237)		(0.2061)
Market–newspaper FE?	X	X	X	X
Zip code demographics?		X	X	X
Zip code X market characteristics?		X	X	X
Zip code FE?			X	
Number of observations	16,043	16,043	16,043	16,043
Number of newspapers	290	290	290	290

[a]The dependent variable is log odds ratio $\ln(S_{zn}) - \ln(1 - S_{zn})$. Standard errors (in parentheses) allow for correlation in the error term across observations for the same newspaper. Zip code demographics are log of total population, log of income per capita, percent of population urban, percent white, percent black, population per square mile, share of houses that are owner occupied, and the share of population aged 25 and over whose highest level of schooling is college, all as of 2000. "Zip code X market characteristics" refers to a vector of these characteristics interacted with their analogue at the level of the newspaper's market. An excluded instrument in the model in the last column is an interaction between zip share donating to Republicans and share of Republican in the newspaper's market in 2004. The first-stage F-statistic on the excluded instrument is 8.79.

7. EVIDENCE ON THE SUPPLY OF SLANT

7.1. *Does Consumer Ideology Affect Slant?*

Consistent with Hypothesis S1, slant is highly related to consumer ideology. Figure 4 plots estimated slant $\hat{y}_n$ against the share voting Republican R_n in the newspaper's market. The graph shows clearly that in more Republican markets, newspapers adopt a more right-wing slant. The first column of Table III shows that in an OLS regression, an increase of 10 percentage points in the share voting Republican translates into an increase in slant of 0.015. This coefficient is highly statistically significant, and variation in consumer preferences explains nearly 20 percent of the variation in slant in this specification.

The relationship between slant and consumer ideology is robust to corrections for possible reverse causality from slant to consumer ideology. The second column of Table III (2SLS (two-stage least squares)) shows that the estimated effect of consumer ideology on slant is similar (though less precise) when we instrument for slant with an estimate of the share of the newspaper's market attending church monthly or more during 1972–1998. This variable has a large effect on a market's political leaning (Glaeser, Ponzetto, and Shapiro (2005)), and our estimates using this instrument are valid if the religiosity of

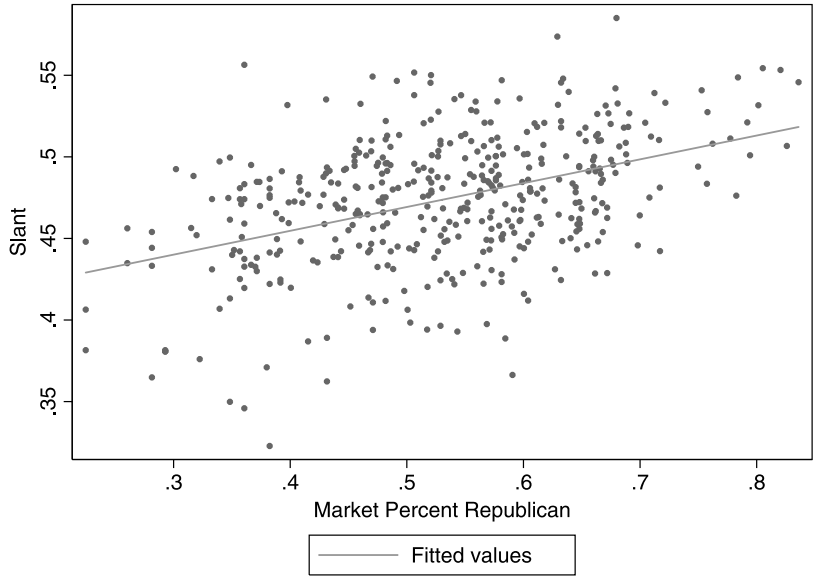

FIGURE 4.—Newspaper slant and consumer ideology. The newspaper slant index against Bush's share of the two-party vote in 2004 in the newspaper's market is shown.

a geographic market is exogenous to the political slant of the market's daily newspaper.

TABLE III

DETERMINANTS OF NEWSPAPER SLANT[a]

	OLS	2SLS	OLS	RE
Share Republican in newspaper's market	0.1460 (0.0148)	0.1605 (0.0612)	0.1603 (0.0191)	0.1717 (0.0157)
Ownership group fixed effects?			X	
State fixed effects?				X
Standard deviation (SD) of ownership effect				0.0062 (0.0037)
Likelihood ratio test that SD of owner effect is zero (p value)				0.1601
Number of observations	429	421	429	429
R^2	0.1859	—	0.4445	—

[a]The dependent variable is slant index ($\hat{y}_n$). Standard errors are given in parentheses. An excluded instrument in the 2SLS model is share attending church monthly or more in the newspaper's market during 1972–1998, which is available for 421 of our 429 observations. The first-stage has coefficient 0.2309 and standard error 0.0450. The RE model was estimated via maximum likelihood. See Section 7.2 for details.

The third column of Table III shows that the estimated effect of consumer ideology is similar when we include fixed effects for ownership groups. This confirms that our result is not driven by a tendency of owners to buy papers in markets where consumers' ideology is similar to their own.

In Gentzkow and Shapiro (2007), we reported a number of additional robustness checks. First, we include controls for several measures of newspaper quality (following Berry and Waldfogel (2003)): the log of the newspaper's number of employees, the log of the number of pages, and the number of Pulitzer prizes from 1970 to 2000. Second, we instrument for consumer ideology with a vector of market demographics predictive of voting: log population, percent black, percent with a college degree, percent urban, and log income per capita. Third, we use a preliminary version of our slant measure for the years 2000 and 2004, along with voting data for both years, to estimate a model with newspaper fixed effects. In all cases, the estimated effect of consumer ideology on slant remains large and statistically significant.

7.2. *Does Ownership Affect Slant?*

Turning to Hypothesis S2, once we account for the propensity of owners to own newspapers in politically and geographically similar markets, we find no evidence that two jointly owned newspapers have a more similar slant than two randomly chosen newspapers. Panel A of Figure 5 plots each newspaper's slant against the average slant of other newspapers with the same owner, revealing a positive and statistically significant correlation. Panel B plots the residual from a regression of slant on the Republican vote share in a paper's market and state fixed effects against the average of this residual among other papers with the same owner. In this panel, there is no visible correlation between the two variables, and the relationship between the variables is no longer significant.

The last column of Table III presents estimates of our preferred supply model—equation (9) under the assumptions of Section 5.2. Our estimate of the variance of the owner effect is small, and we cannot reject the null hypothesis that the variance of the owner effect is zero.

We find no evidence that slant is related to owner ideology, as proxied by political donations. In Figure 6, we plot the relationship between slant and the share of contributions going to Republican candidates for three categories of contributions: (i) those from executives at firms that own multiple U.S. newspapers, (ii) those from executives at independent newspapers (not jointly owned with any other U.S. paper), and (iii) corporate contributions by newspaper firms. The correlation between slant and contributions is weak and statistically insignificant. This remains true in regressions that control for the percent voting Republican in each paper's market (see online Appendix B, Table B.II). Taking donations as a proxy for owner ideology, then, we do not find evidence for Hypothesis S2.

In Gentzkow and Shapiro (2007), we reported additional evidence on the role of ownership in determining slant. We show in a range of random effects

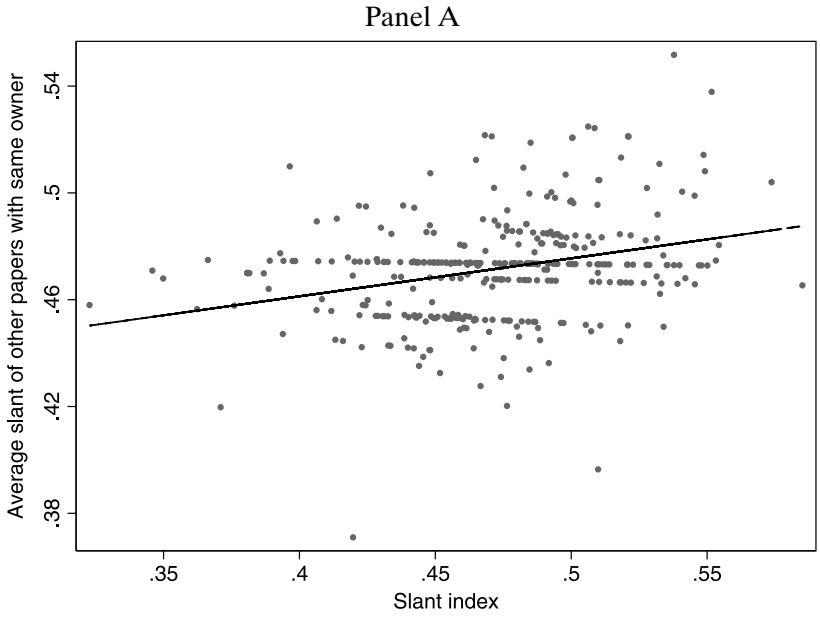

FIGURE 5.—Newspaper slant and ownership. Panel A shows average slant of co-owned newspapers graphed against a newspaper's own slant (correlation = 0.29, $p < 0.001$). Panel B parallels Panel A, but measures slant using residuals from a regression of slant on percent Republican in market and dummies for the state in which the newspaper is located (correlation = 0.09, $p = 0.11$).

models that the owner effect diminishes as we control more tightly for geography, and that it is largely eliminated by controlling for the Republican vote share and Census division fixed effects. In contrast, the role of consumer characteristics grows stronger as we focus on variation in slant within geographic areas. We also examine three important ownership changes that occur during a period (2000–2005) for which we have computed a preliminary slant index. We find no clear evidence that acquired newspapers' slant moves closer to the mean slant of newspapers in the acquiring group.

8. IMPLICATIONS OF THE MODEL

Table IV presents a series of calculations that expose the model's economic implications.

The first row of Table IV presents the observed slant of the average newspaper in the sample. The second row of Table IV presents the profit-maximizing slant of the average newspaper in the sample. Though statistically distinguishable, the two are close in magnitude. At our point estimate, the average news-

Panel B

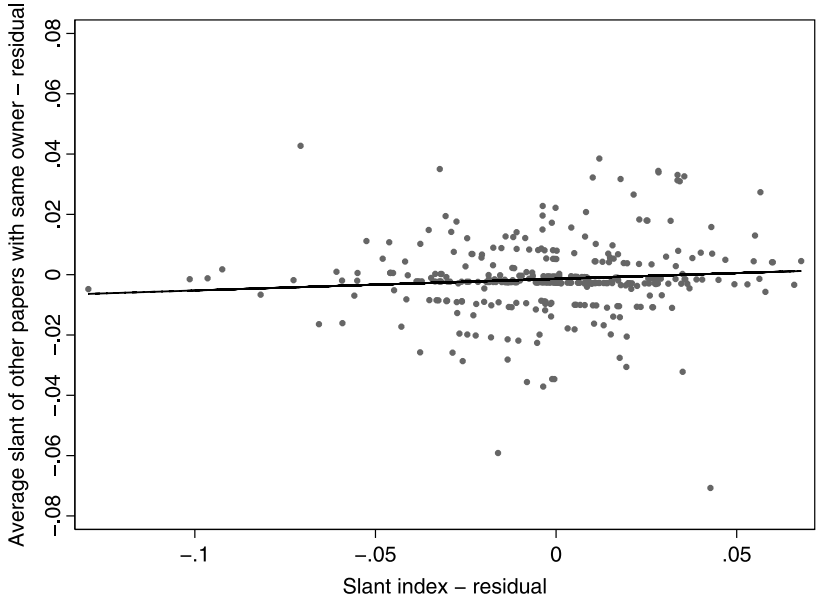

FIGURE 5.—(Continued.)

paper would move slightly to the left in a counterfactual world in which all newspapers choose exactly the profit-maximizing value of slant.

Newspapers could deviate systematically from profit maximization on average due to owner ideology (Balan, DeGraba, and Wickelgren (2009)), pressure from incumbent politicians (Besley and Prat (2006)), or the tastes of reporters (Baron (2006)). A large popular literature has argued that such forces create an overall conservative (Alterman (2003), Franken (2003)) or liberal (Coulter (2003), Goldberg (2003)) bias in the media. Our data do not show evidence of an economically significant bias relative to the benchmark of profit maximization.

The third row of Table IV presents the percent loss in circulation that the average newspaper would experience if it were to deviate by 1 standard deviation from the profit-maximizing level of slant. We estimate an economically large effect of about 18 percent, though the precision of this estimate is limited.

The last two rows of Table IV present the shares of the within-state variation in slant that can be explained by variation in consumer and owner ideology, respectively. At our point estimates, consumer ideology explains 22 percent of the within-state variation in slant, while owner ideology explains only 4 percent. Put differently, our point estimates imply that eliminating cross-market

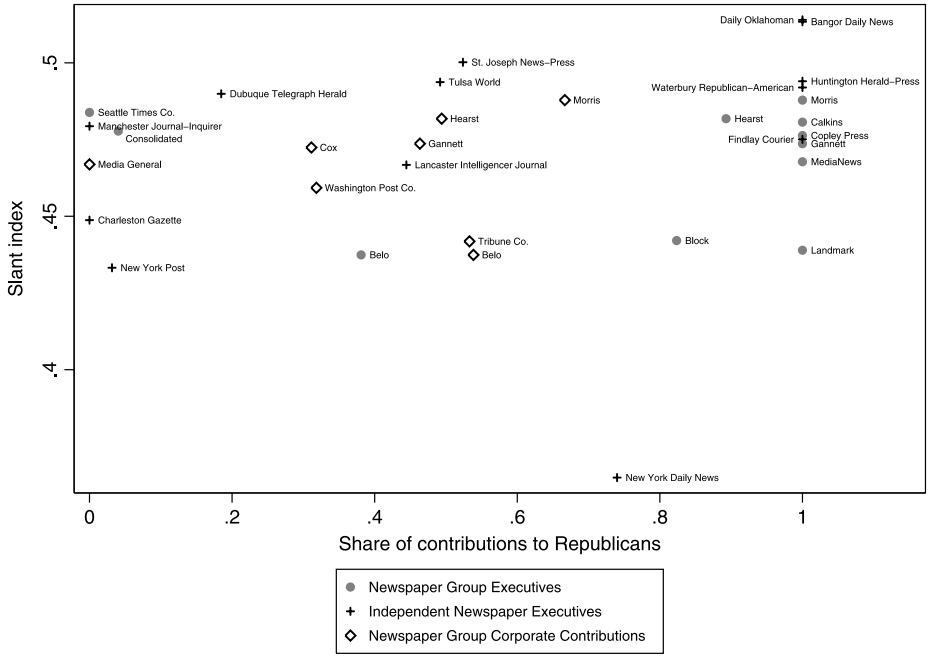

FIGURE 6.—Newspaper slant and political contributions. The average slant of newspapers owned by a firm is graphed against the share of total dollars going to Republicans within each category of contributions. Correlation coefficients are -0.04 ($p = 0.90$) for newspaper group executives, 0.29 ($p = 0.34$) for independent newspaper executives, and 0.01 ($p = 0.97$) for newspaper group corporate contributions.

TABLE IV

ECONOMIC INTERPRETATION OF MODEL PARAMETERS[a]

Quantity	Estimate
Actual slant of average newspaper	0.4734
	(0.0020)
Profit-maximizing slant of average newspaper	0.4600
	(0.0047)
Percent loss in variable profit to average newspaper	0.1809
from moving 1 SD away from profit-maximizing slant	(0.1025)
Share of within-state variance in slant from consumer ideology	0.2226
	(0.0406)
Share of within-state variance in slant from owner ideology	0.0380
	(0.0458)

[a]Standard errors, given in parentheses, are from the delta method. The sample in the fist three rows includes 290 newspapers in the demand sample. The sample in the last two rows includes 429 newspapers in the supply sample. The calculation in the fourth row is $(\hat{\lambda}_1^s)^2$ times the within-state variance in R_n, divided by the within-state variance of $\hat{y}_n$. The calculation in the last row is $\hat{\sigma}_\mu^2$ divided by the within-state variance of $\hat{y}_n$.

diversity in consumer ideology would reduce the variance of measured slant by 22 percent, whereas eliminating heterogeneity in owner ideology (say by having all newspapers jointly owned) would reduce it by only 4 percent. We can reject the hypothesis that the share of variance explained by consumers and owners is the same ($p = 0.003$).

9. OTHER DETERMINANTS OF SLANT

We have interpreted the observed relationship between slant and consumer ideology as evidence that newspapers cater to their readers. Here, we consider two alternative explanations:

(i) Incumbent politicians influence news content (Besley and Prat (2006)), and incumbent politicians' ideology is correlated with consumer ideology.

(ii) Reporters and editors are drawn from the local population, have ideologies correlated with those of local consumers, and are willing to sacrifice wage income to represent their own views in the newspaper (Baron (2006)).

Ideology of Incumbent Politicians

If incumbent politicians influence news content, then any correlation between incumbent politicians' ideology and consumer ideology could bias our results. In regression models reported in online Appendix B, we find no evidence that slant is related to the party affiliation of local elected officials. Controlling for consumer ideology, having a Republican governor (as of the end of 2005) is associated with a statistically insignificant leftward shift in slant of about 0.9 percentage points, with a confidence interval that rules out a rightward shift larger than about 0.5 percentage points (1/8 of a standard deviation). We also find that, controlling for consumer ideology, the Republican share of representatives to the U.S. House from districts in the newspaper's market (as of the 109th Congress) has a statistically insignificant negative effect on slant. The coefficient implies that moving from a completely Democratic to a completely Republican delegation reduces newspaper slant by 0.004, with a confidence interval that excludes substantial positive effects.

Ideology of Local Reporters and Editors

If local reporters/editors always had the same ideological preferences as consumers, a model where slant responds only to consumers and a model where it also responds to reporters/editors would be observationally equivalent. The important economic question is therefore how slant would be chosen in the event that reporters' and editors' ideologies diverged from those of consumers. For a number of reasons, we believe that it is unlikely that reporter/editor ideology would exert a significant influence in such a case.

Consider a case where consumers' preferred slant is 1 standard deviation to the right of that of local editors and reporters. The local newspaper considers

whether to choose reporters' or consumers' preferred slant. (For simplicity, suppose this choice is either/or.) The cost of satisfying consumer demand is that the newspaper must pay more to bring in qualified reporters and editors from elsewhere and possibly train them in local knowledge, or convince local staff to deviate from their personal ideologies. According to our demand estimates, the benefit is an increase of 18 percent in variable profits. A crude estimate is that the salaries of editors and reporters are on the order of 10 percent of variable profits for a typical newspaper.[9] Therefore, for reporters' tastes to overwhelm consumer demand, equally qualified reporters willing to report as consumers wish would need to cost 18 percent/10 percent = 180 percent more than those drawn from the local population.

That the cost of qualified reporters could be so high seems especially unlikely given that the market for editors and reporters is not highly localized. In a regression model using Census microdata, we find that reporters and editors are 8 percentage points *more* likely than other professionals to live in a state other than the one in which they were born, controlling for education, age, gender, and race.[10] These "outside" reporters and editors are not of lower quality: reporters and editors born outside their current state of residence earn, if anything, somewhat *more* than those working in their states of nativity. Survey data also show that the average college-educated journalist has nearly a 40 percent chance of working in a Census division other than the one in which he or she attended college (Weaver and Wilhoit (1996)), considerably higher than the average among other college-educated workers.[11]

Put differently, the elasticity of reporters and editors of different types into a given local market is likely to be very high, as each market draws from the same large national pool of talent. Given consumers' strong demand for like-minded slant, if the tastes of local readers and potential local reporters varied independently, we would expect the tastes of readers to dominate in the determination of equilibrium slant.

As a separate test of the influence of local reporters' ideology, we have constructed a version of our slant measure using only stories written by newspa-

[9]Gentzkow (2007) estimated that the *Washington Post*'s variable profit per *daily* copy sold was $1.83 in 2004. Applying the same profit rate to Sunday copies (probably an understatement) gives a total yearly variable profit of $539 million. *Burrelle's/Luce Media Directory 2001* (Burrelle's Information Services (2001)) lists 222 reporters and 175 editors working for the *Post*. If we assume that the average reporter's salary is $90,000 per year and the average editor's salary is $125,000 per year, we estimate the *Post*'s wage bill for reporters and editors to be about $42 million per year, or about 8 percent of variable profits.

[10]They are also three percentage points more likely to have moved in the past five years. These figures are coefficients on reporter/editor dummies in regressions using data from the 1980, 1990, and 2000 Censuses (Ruggles et al. (2004)). The sample is restricted to 25- to 55-year-old workers in professional occupations (1950 occupation codes 000–099). Wage regressions reported below are restricted to prime-age male reporters and editors working full time.

[11]We are extremely grateful to Lisa Kahn for providing the appropriate calculations from the 1979 National Longitudinal Study of Youth (NLSY).

pers' Washington DC bureaus. The reporters and editors of these stories typically live and work in Washington and not in their newspapers' home markets. If slant were determined largely by the geographic home of the editorial staff, we would expect much more homogeneous slant in Washington bureau stories than in locally written stories. In fact, a regression of the slant of Washington bureau stories on consumer ideology yields a positive and statistically significant coefficient, with a value not statistically distinguishable from the coefficient we obtain when we use the overall slant measure. (We note, however, that many papers do not have Washington bureaus, which limits the statistical power of this test.)

Note that the preceding argument is fully consistent with an equilibrium correlation between consumers' and reporters' ideologies; indeed, we would expect such a correlation if reporters have a comparative advantage in writing with a slant consistent with their own views. While we do not have direct evidence on the institutional mechanism through which newspapers "choose" their slant, the choice of editorial staff (along with choice of topics and explicit style policies) seems like a plausible channel through which newspaper content is calibrated to the views of the local population.

10. CONCLUSIONS

In this paper, we develop and estimate a new measure of slant that compares the use of partisan language in newspapers with that of Democrats and Republicans in Congress. Our measure is computable with a minimum of subjective input, is related to readers' subjective ratings of newspaper slant, and is available for newspapers representing over 70 percent of the daily circulation in the United States.

Combining our measure with zip code-level circulation data, we show that consumer demand responds strongly to the fit between a newspaper's slant and the ideology of potential readers, implying an economic incentive for newspapers to tailor their slant to the ideological predispositions of consumers. We document such an effect and show that variation in consumer preferences accounts for roughly one-fifth of the variation in measured slant in our sample.

By contrast, we find much less evidence for a role of newspaper owners in determining slant. While slant is somewhat correlated across co-owned papers, this effect is driven by the geographic clustering of ownership groups. After controlling for the geographic location of newspapers, we find no evidence that the variation in slant has an owner-specific component. We also find no evidence that pressure from incumbent politicians or the tastes of reporters are important drivers of slant.

Taken together, our findings suggest that ownership diversity may not be a critical precondition for ideological diversity in the media, at least along the dimension we consider. This conclusion has broad implications for the regulation of ownership in the media.

We wish to stress three important caveats, however.

First, our measure of slant is a broad aggregate that includes coverage of many different topics over a reasonably long window of time. Owners, politicians, or reporters may still exert significant influence on coverage of specific domains in which their interests are especially strong. For example, Gilens and Hertzman (2009) showed that the 1996 Telecommunications Act received more favorable coverage from newspapers whose parent companies stood to gain from the act's passage. In such areas, where the financial interest of the owner is strong relative to the likely interest of the reader, it is not surprising to see an important effect of ownership, even in light of our finding that ownership is not predictive of our broad index of slant.

Second, our results may not extend to settings with significantly different legal or institutional environments—less developed markets, more state ownership, less freedom of the press. Silvio Berlusconi's influence on Italian media is a case in point (Anderson and McLaren (2009), Durante and Knight (2009)).

Finally, finding that ownership is not an important driver of content diversity does not imply that the market produces the *optimal* level of diversity. In particular, it remains true that virtually all local newspaper markets are monopolies, and the number of independent sources for local news is many cities is correspondingly small. How diversity and welfare are affected by the degree of local newspaper competition remains an important area for future research.

APPENDIX A: DETAILS ON NEWS SEARCHES

A.1. *Mechanics of* Congressional Record

We use an automated script to download the *Congressional Record* from thomas.loc.gov. Our data base of *Congressional Record* text is incomplete, mostly due to errors in the website that archives the *Congressional Record*. These errors affect a relatively small share of documents in the *Congressional Record* (roughly 15 percent).

We apply a second script to the downloaded text to ascertain the speaker of each passage. We wish to focus on floor speeches rather than text that is primarily procedural, so we exclude speech by officers such as the Clerk, the Speaker of the House, and the President of the Senate. We also exclude block quotations, text that is inserted into the *Record* from other sources such as reports or letters, and nonspeech items like records of roll-call votes.

Before producing phrase counts, we remove extremely common words ("stopwords"). We use the list from Fox (1990), augmented with a list of proper nouns that appear frequently in procedural text—days of the week, the Hart

Senate Office Building, and the Dirksen Senate Office Building. We also exclude the names of major newspapers.

We use the Porter Stemmer (tartarus.org/martin/PorterStemmer/) to strip words down to their linguistic roots. This means that phrases in the *Congressional Record* that differ only in either stopwords or suffixes are equivalent in our algorithm. For example, "war on terror," "war against terror," and "wars on terror" would all appear in the preprocessed *Congressional Record* as "war terror" and thus be treated as the same phrase.

A.2. *Mechanics of Newspaper Searches*

Following the steps outlined in Section 3.1, we identify 1000 phrases to use in our analysis. We wish to count the number of times each of these 1000 phrases appears in each of our sample of newspapers using the ProQuest and NewsLibrary data bases.

Among our 433 newspapers, data are available for 394 from NewsLibrary and for 164 from ProQuest, with an overlap of 125 newspapers. Among the newspapers that overlap between the two data bases, the correlation between the counts for our 1000 phrases is 0.85. In cases of overlap, we use the NewsLibrary counts for analysis.

The two data bases do not agree perfectly for several reasons, including differences in the set of articles newspapers choose to post to each data base and differences in how the two data bases permit us to identify editorials and opinion pieces (see below). An important third reason is that the data bases are dynamic: content is added over time, so that searches conducted at different times may produce different results. As a consequence, one potential source of disagreement between ProQuest and NewsLibrary is a difference in the posting lag between the two data bases.

Because of the preprocessing steps above (stopword removal and stemming), each of our 1000 phrases thus corresponds to a group of one, two, or several original phrases, and it is these original phrases that we search for in the data bases.

The set of original phrases we search is slightly restricted for two reasons. First, the ProQuest data base limits search strings to 75 characters. We therefore drop any original phrase longer than 75 characters. Second, our data base of *Congressional Record* text has improved over time as we have adjusted for errors in the source website and improved our parsing algorithm. The set of original phrases included in each group is based on a slightly older version of the *Congressional Record* text than the one used for our main analysis, so it omits some relatively rare original phrases.

We search for each group of original phrases (connected with the OR operator) in the All Text field (NewsLibrary) or Document Text field (ProQuest), restricted to 2005 and with the following terms excluded from the Headline and Author fields: "editor," "editorial," "associated press," "ap," "opinion," "op-ed," and "letter."

A.3. *Audit Study*

Our searches are designed to isolate the slant of news content produced independently by each paper. The way stories are archived and classified in the data bases means that we can only imperfectly separate these stories from other kinds of content such as opinion pieces and wire stories. To provide a more precise picture of the kinds of content we are measuring, we have audited the results for seven phrases chosen from Table I. For each phrase, we looked at the full set of hits for the papers included in the NewsLibrary data base and recorded whether they appeared to be (i) independently produced news stories, (ii) AP wire stories, (iii) other wire stories, (iv) letters to the editor, or (v) opinion pieces (including unsigned editorials). Because we do not have access to the full text of articles in NewsLibrary, this classification is based on the headline and first paragraph of the story.

In a separate exercise, we use results from the papers we can search in the ProQuest data base (for which we can retrieve full text articles) to record the number of times each phrase appears in quotation.

The results are shown in Table A.I. Overall, approximately 71 percent of our hits are independently produced news stories. Of the remainder, 22 percent are either clearly or possibly opinion, 3 percent are letters to the editor, and 3 percent are wire stories. The table also shows that these shares are heterogeneous across phrases. For example, the share of opinion pieces ranges from 12 percent for "global war on terrorism" to 51 percent for "death tax." The results also show that only 10 percent of our hits appear in quotations, with the share ranging from 3 percent for "child support enforcement" to 36 percent for "death tax." We have also spot checked the articles that are being excluded from our search results and verified that virtually all of them are, as desired, either wire stories or opinion pieces.

As a final check, we have also computed the share of phrases appearing in direct quotes of local congresspeople, which could cause a mechanical correlation between slant and the political leanings of local markets. Among 10 randomly chosen papers (representing different levels of circulation), we hand coded the frequency of uses of the top 50 phrases in direct quotes of congresspeople. On average, such quotes account for only 0.3 percent of the phrase hits in this sample.

Taken together, the results confirm that our measure is primarily picking up the slant of independently produced news stories, with some weight given to opinion pieces.

REFERENCES

AAS, K., AND L. EIKVIL (1999): "Text Categorisation: A Survey," Report 941, Norwegian Computing Center. [38]

AKHAVAN-MAJID, R., A. RIFE, AND S. GOPINATH (1991): "Chain Ownership and Editorial Independence: A Case Study of Gannett Newspapers," *Journalism and Mass Communication Quarterly*, 68, 59–66. [38]

TABLE A.I

AUDIT OF SEARCH RESULTS[a]

Phrase	Total Hits	Share of Hits in Quotes	Share of Hits That Are					
			AP Wire Stories	Other Wire Stories	Letters to the Editor	Maybe Opinion	Clearly Opinion	Independently Produced News
Global war on terrorism	2064	16%	3%	4%	1%	2%	10%	80%
Malpractice insurance	2190	5%	0%	0%	1%	3%	12%	84%
Universal health care	1523	9%	1%	0%	7%	8%	28%	56%
Assault weapons	1411	9%	3%	12%	4%	1%	25%	56%
Child support enforcement	1054	3%	0%	0%	1%	2%	11%	86%
Public broadcasting	3375	8%	1%	0%	2%	4%	22%	71%
Death tax	595	36%	0%	0%	2%	5%	46%	47%
Average (hit weighted)		10%	1%	2%	3%	3%	19%	71%

[a]Authors' calculations based on ProQuest and NewsLibrary data base searches. See Appendix A for details.

ALTERMAN, E. (2003): *What Liberal Media? The Truth About Bias and the News*. Basic Books. [60]
ANDERSON, S. P., AND J. MCLAREN (2009): "Media Mergers and Media Bias With Rational Consumers," Working Paper, University of Virginia. [65]
ANSOLABEHERE, S., R. LESSEM, AND J. M. SNYDER JR. (2006): "The Orientation of Newspaper Endorsements in U.S. Elections, 1940–2002," *Quarterly Journal of Political Science*, 1, 393–404. [38]
ANTWEILER, W., AND M. Z. FRANK (2004): "Is All That Talk Just Noise? The Information Content of Internet Message Boards," *Journal of Finance*, 59, 1259–1294. [38]
BAGDIKIAN, B. H. (2000): *The Media Monopoly* (Sixth Ed.). Boston: Beacon Press. [36]
BALAN, D. J., P. DEGRABA, AND A. L. WICKELGREN (2009): "Ideological Persuasion in the Media," Mimeo, Federal Trade Commission. [38,60]
BARON, D. P. (2006): "Persistent Media Bias," *Journal of Public Economics*, 90, 1–36. [37,38,60,62]
BERRY, S., AND J. WALDFOGEL (2003): "Product Quality and Market Size," Working Paper 9675, NBER. [58]
BESLEY, T., AND A. PRAT (2006): "Handcuffs for the Grabbing Hand? Media Capture and Government Accountability," *American Economic Review*, 96, 720–736. [37,38,60,62]
BURRELLE'S INFORMATION SERVICES (2001): *Burrelle's/Luce Media Directory—2001 Edition*. Livingston, NJ: Burrelle's Information Services. [63]
COMMISSION ON FREEDOM OF THE PRESS (1947): *A Free and Responsible Press: A General Report on Mass Communication: Newspapers, Radio, Motion Pictures, Magazines, and Books*. Chicago, IL: The University of Chicago Press. [36]
COULTER, A. (2003): *Slander: Liberal Lies About the American Right*. New York: Three Rivers Press. [60]
DELLAVIGNA, S., AND E. KAPLAN (2007): "The Fox News Effect: Media Bias and Voting," *Quarterly Journal of Economics*, 122, 1187–1234. [36]
DRANOVE, D., A. GRON, AND M. J. MAZZEO (2003): "Differentiation and Competition in HMO Markets," *Journal of Industrial Economics*, 51, 433–454. [38]
DUBÉ, J.-P., G. J. HITSCH, AND P. MANCHANDA (2005): "An Empirical Model of Advertising Dynamics," *Quantitative Marketing and Economics*, 3, 107–144. [38]
DURANTE, R., AND B. KNIGHT (2009): "Partisan Control, Media Bias, and Viewer Responses: Evidence From Berlusconi's Italy," Working Paper 14762, NBER. [65]
EINAV, L. (2007): "Seasonality in the US Motion Picture Industry," *RAND Journal of Economics*, 38, 127–145. [38]
FEDERAL COMMUNICATIONS COMMISSION (2003): *Report and Order and Notice of Proposed Rulemaking*. Washington, DC: Federal Communications Commission. [36]
FOX, C. (1990): "A Stop List for General Text," *SIGIR FORUM*, 24, 19–35. [65]
FRANKEN, A. (2003): *Lies and the Lying Liars Who Tell Them: A Fair and Balanced Look at the Right*. Boston, MA: E. P. Dutton. [60]
GENTZKOW, M. (2006): "Television and Voter Turnout," *Quarterly Journal of Economics*, 121, 931–972. [36]
——— (2007): "Valuing New Goods in a Model With Complementarity: Online Newspapers," *American Economic Review*, 97, 713–744. [41,63]
GENTZKOW, M. A., AND J. M. SHAPIRO (2004): "Media, Education, and Anti-Americanism in the Muslim World," *Journal of Economic Perspectives*, 18, 117–133. [36]
——— (2006): "Media Bias and Reputation," *Journal of Political Economy*, 114, 280–316. [37,38]
——— (2007): "What Drives Media Slant? Evidence From U.S. Daily Newspapers," Working Paper 12707, NBER. [49,50,58]
——— (2010): "Supplement to 'What Drives Media Slant? Evidence From U.S. Daily Newspapers'," *Econometrica Supplemental Material*, 78, http://www.econometricsociety.org/ecta/Supmat/7195_tables-figures.pdf. [39]
GENTZKOW, M. A., E. L. GLAESER, AND C. D. 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, ed. by E. L. Glaeser and C. Goldin. Chicago, IL: University of Chicago Press, 187–230, Chapter 6. [38]

GEORGE, L. (2007): "What's Fit to Print: The Effect of Ownership Concentration on Product Variety in Newspaper Markets," *Information Economics and Policy*, 19, 285–303. [38]

GEORGE, L., AND J. WALDFOGEL (2003): "Who Affects Whom in Daily Newspaper Markets?" *Journal of Political Economy*, 111, 765–784. [37,51,52]

GERBER, A. S., D. KARLAN, AND D. BERGAN (2009): "Does the Media Matter? A Field Experiment Measuring the Effect of Newspapers on Voting Behavior and Political Opinions," *American Economic Journal: Applied Economics*, 1 (2), 35–52. [36]

GILENS, M., AND C. HERTZMAN (2009): "Corporate Ownership and News Bias: Newspaper Coverage of the 1996 Telecommunications Act," *Journal of Politics*, 62, 369–386. [65]

GIMPEL, J. G., F. E. LEE, AND J. KAMINSKI (2006): "The Political Geography of Campaign Contributions in American Politics," *Journal of Politics*, 68, 626–639. [41]

GLAESER, E. L., G. A. M. PONZETTO, AND J. M. SHAPIRO (2005): "Strategic Extremism: Why Republicans and Democrats Divide on Religious Values," *Quarterly Journal of Economics*, 120, 1283–1330. [56]

GLASSER, T. L., D. S. ALLEN, AND S. E. BLANKS (1989): "The Influence of Chain Ownership on News Play: A Case Study," *Journalism Quarterly*, 66, 607–614. [36]

GOLDBERG, B. (2003): *Bias: A CBS Insider Exposes How the Media Distort the News*. New York: Perennial. [60]

GRAETZ, M. J., AND I. SHAPIRO (2005): *Death by a Thousand Cuts: The Fight Over Taxing Inherited Wealth*. Princeton, NJ: Princeton University Press. [42]

GROSECLOSE, T., AND J. MILYO (2005): "A Measure of Media Bias," *Quarterly Journal of Economics*, 120, 1191–1237. [36,38,42]

HAMILTON, J. T. (2004): *All the News That's Fit to Sell: How the Market Transforms Information Into News*. Princeton, NJ: Princeton University Press. [38]

HARRIS INTERACTIVE (2006): "Seven in 10 U.S. Adults Say They Watch Broadcast News at Least Several Times a Week," *The Harris Poll*, 20. [36]

LACY, S., AND T. F. SIMON (1997): "Intercounty Group Ownership of Daily Newspapers and the Decline of Competition for Readers," *Journalism and Mass Communication Quarterly*, 74, 814–825. [53]

LARCINESE, V., R. PUGLISI, AND J. M. SNYDER (2007): "Partisan Bias in Economic News: Evidence on the Agenda-Setting Behavior of U.S. Newspapers," Working Paper 13378, NBER. [38]

LAVER, M., K. BENOIT, AND J. GARRY (2003): "Extracting Policy Positions From Political Texts Using Words as Data," *American Political Science Review*, 97, 311–331. [38]

LUNTZ, F. (2005): *Learning From 2004 … Winning in 2006*. Washington, DC: Luntz Research Companies. [44,45]

MARTIN, H. J. (2003): "Some Effects From Horizontal Integration of Daily Newspapers on Markets, Prices, and Competition," in *Proceedings of the Annual Meeting of the Association Foe Education in Journalism and Mass Communication, Media Management and Economics Division*, Kansas City, MO, July 30–August 2, 2003. [53]

MAZZEO, M. J. (2002a): "Product Choice and Oligopoly Market Structure," *RAND Journal of Economics*, 33, 221–242. [38]

——— (2002b): "Competitive Outcomes in Product-Differentiated Oligopoly," *Review of Economics and Statistics*, 84, 716–728. [38]

MULLAINATHAN, S., AND A. SHLEIFER (2005): "The Market for News," *American Economic Review*, 95, 1031–1053. [37,38,48]

MYERS, C. K. (2008): "Discrimination as a Competitive Device: The Case of Local Television News," *The B.E. Journal of Economic Analysis and Policy*, 8, Article 28. [38]

NEWSPAPER ASSOCIATION OF AMERICA (2006): "The Source: Newspapers by the Numbers." [36]

PRITCHARD, D. (2002): "Viewpoint Diversity in Cross-Owned Newspapers and Television Stations: A Study of News Coverage of the 2000 Presidential Campaign," Working Paper, FCC Media Ownership Working Group. [36]

PUGLISI, R. (2008): "Being the New York Times: The Political Behavior of a Newspaper," Mimeo, ECARES-ULB. [38]

PUTNAM, R. D. (2000): *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon and Schuster. [40]

RUGGLES, S., M. SOBEK, T. ALEXANDER, C. A. FITCH, R. GOEKEN, P. K. HALL, M. KING, AND C. RONNANDER (2004): *Integrated Public Use Microdata Series: Version 3.0*. Minneapolis, MN: Minnesota Population Center. Available at http://www.ipums.org. [63]

SEIM, K. (2006): "An Empirical Model of Firm Entry With Endogenous Product-Type Choices," *RAND Journal of Economics*, 37, 619–640. [38]

STEVENSON, R. W. (2005): "President Makes It Clear: Phrase Is 'War on Terror'," *New York Times*, August 4. [45]

STRÖMBERG, D. (2004): "Radio's Impact on Public Spending," *Quarterly Journal of Economics*, 119, 189–221. [36]

SUEN, W. (2004): "The Self-Perpetuation of Biased Beliefs," *Economic Journal*, 114, 377–396. [38]

SWEETING, A. (2007): "Dynamic Product Repositioning in Differentiated Product Industries: The Case of Format Switching in the Commercial Radio Industry," Working Paper 13522, NBER. [38]

——— (2008): "The Effects of Horizontal Mergers on Product Positioning: Evidence From the Music Radio Industry," Mimeo, Duke University. [38]

U.S. SUPREME COURT (1945): "Associated Press v. United States," Washington, DC. [35]

WEAVER, D. H., AND G. C. WILHOIT (1996): *The American Journalist in the 1990s: U.S. News People at the End of an Era*. Mahwah, NJ: Lawrence Erlbaum Associates. [63]

University of Chicago Booth School of Business, 5807 S. Woodlawn Avenue, Chicago, IL 60637, U.S.A. and NBER; gentzkow@chicagobooth.edu

and

University of Chicago Booth School of Business, 5807 S. Woodlawn Avenue, Chicago, IL 60637, U.S.A. and NBER; jesse.shapiro@chicagobooth.edu.

Manuscript received May, 2007; final revision received August, 2009.