

## Jesse Shapiro

**Editor's Note:** *This is an abbreviated version of EF's conversation with Jesse Shapiro. For additional content, go to our website: [www.richmondfed.org/publications](http://www.richmondfed.org/publications)*

The objectivity of news and the crucial role of media in democracy have rarely been hotter topics than they are today. News reporting inherently involves selection, notes Brown University's Jesse Shapiro: There is no one correct way to present the same set of events. And with that process of selection may come charges of bias.

But what determines the extent of media bias and how it leans? Shapiro emphasizes the role of market forces in work with frequent co-author Matthew Gentzkow of Stanford University and others. By training machines to classify newspaper text, he has found that newspaper ownership and political incumbents do not influence media bias as much as conventional wisdom would suggest — the demand of consumers for “like-minded” news plays a much bigger role. Using historical data from the days when many cities had multiple local newspapers, he also found that competitive news markets promote ideological diversity. As for the ultimate competitive marketplace for news — the Internet — he found using browsing histories that online news consumption is far less politically polarized than many people have feared.

Shapiro's work doesn't stop at media markets but often similarly relies on novel sources and enormous datasets to address other topics. It is no secret that Republicans and Democrats often speak different languages; Shapiro and co-authors looked at nearly 150 years of congressional speech to identify just when — and potentially why — politically polarized language came to be. And looking at more than 6 billion retail purchases, his work found that the way households use food stamps may pose a challenge to basic economic theory.

Shapiro notes the potential power of social media as a tool for studying how we think and interpret news events in real time — his next focus of study.

Before becoming the George S. and Nancy B. Parker Professor of Economics at Brown University in 2015, Shapiro was a professor at the University of Chicago Booth School of Business. He also is a research associate at the National Bureau of Economic Research, where he is a member of the labor studies, political economy, and industrial organization programs. Renee Haltom interviewed him at his office at Brown in May 2017.



**EF:** You and Matt Gentzkow found that consumer tastes can be a stronger source of media bias than other factors people often suspect, such as who owns the news source. How did you tackle that question?

**Shapiro:** We did text analysis of newspapers using methods borrowed from machine learning. These methods go back at least to the 1960s to authorship detection, like figuring out whether Shakespeare really wrote a given play or who wrote the *Federalist Papers*. That works extremely well, but it tends to rely on having what's called a training set: You have some documents where you know who wrote them and use them to find the tell for any given author — what are the things they say that other people don't tend to say, how they construct sentences, or even how they use commas. Then you look for the features in other documents where you don't know the author to get a sense of who's most likely to have written them.

What we were trying to figure out is which newspapers are right-leaning and which newspapers are left-leaning and by how much. In the context of the news media in the United States, there isn't really a training set. So we took an idea that was developed by Tim Groseclose and Jeffrey Milyo to use the *Congressional Record* as the training set. We have a lot of text by speakers who have a known political affiliation — what party they belong to and how they vote on issues. Then we find the phrases that are diagnostic of the speaker's party. We came up with things like

“death tax” for Republicans and “estate tax” for Democrats, or “personal retirement accounts” for Republicans and “private retirement accounts” for Democrats, or “the war in Iraq” for Democrats and “the war on terror” for Republicans. We could then look for those key-

words or key phrases in newspapers and answer the question: If this newspaper were a speaker in Congress, would it be more likely to be affiliated with the Republican Party or the Democratic Party? That’s our quantitative answer to how right-leaning or left-leaning a newspaper is.

There are a few advantages to taking that kind of approach, as opposed to having a research assistant read newspaper articles and classify them. One advantage is that there’s a portability to the method: If you have some parliamentary text and newspapers, then you can apply the same method to other contexts, and people have. It’s also very scalable. You can have research assistants read a few thousand newspaper articles but not all of them. Whereas previous work had used samples of maybe a couple dozen news outlets, we were able to get up to more than 400. You lose some of the fidelity of having a human reading everything, but the gain in scale means that you can ask questions that are otherwise very difficult to ask, such as: How is the political position of the newspaper related to the position of its readership? How is the position of the newspaper related to the position of its owners? And how do those two things trade off in determining how newspapers cover the news?

What we found is that newspapers with a more Republican customer base are much more Republican than newspapers in more Democratic markets. And once you control for geography, there’s very little evidence of an influence of owner ideology — whether you measure that by the positions of the other newspapers owned by that owner or by the owner’s donations to different political parties. There really isn’t much evidence that the owner plays a big role in how a newspaper slants the news.

**EF: What conditions are likely to make the drive to cater to owners stronger?**

**Shapiro:** I think the drive to cater to *consumers* is probably stronger when you have a robust commercial marketplace for news media, like we generally have in the United States. Obviously, some countries don’t have a very robust commercial market, and newspaper owners have strong ties to government, or are in government, and there the balance of incentives is different.

Even in the United States, if you look at episodes where there’s evidence that newspapers have been tilted in a way that’s biased toward the owner, it seems to line up with

---

There is something to be said for removing the façade of objectivity [in media] and recognizing that journalists are people and that the journalistic process is a human process.

---

what I think economics would predict. There’s a nice paper that looks at coverage of the 1996 Telecommunications Act, which was a bill that went before Congress that really impacted the bottom line of media companies. It was about a bunch of things, including ownership

rules for television. There’s evidence that the slant of a newspaper’s coverage at that time was very correlated with the incentives of the entities that owned them. To me that makes a lot of sense, because this is an issue where consumers don’t have very strong opinions: They’re not likely to be turned off by an editorial saying that the Telecommunications Act should or should not be passed. So even in the United States, when that happens, you can see a tilt in the reporting.

**EF: Nowadays, objectivity is an explicit professional standard in much news reporting. This wasn’t always the case — it used to be common for newspapers to declare party affiliation, for example. Are there conditions under which media bias can be productive?**

**Shapiro:** That’s a tough question. After a lot of time working on this, unfortunately, I still don’t feel like I totally understand why readers have a demand for like-minded news. There are theories in the literature that it’s because people have a psychological bias and want to see their beliefs confirmed, and there are theories that say it’s a kind of rational inference about credibility — Matt and I have a paper that takes that view — and the evidence doesn’t adjudicate very well among these different explanations. So I think there are some things to learn about that.

I don’t know that I would advocate for going back to newspapers declaring party preferences, but I think it has the advantage that when you’ve declared an affiliation, your cards are on the table and you have more license to say what you really think. Take reporting on scientific issues like climate change, the subject of one of my recent papers. Believing in anthropogenic climate change is seen as a position aligned with the left side of the political spectrum, even though it’s really a scientific question. If you had papers that were explicitly right wing and they said, “Actually, the evidence is that there is anthropogenic climate change,” that would carry a lot of weight.

There is something to be said for removing the façade of objectivity and recognizing that journalists are people and that the journalistic process is a human process. It involves perspective and embracing that a little more as opposed to shying away from that and trying to conceal it and in so doing throwing away information that might be really valuable to consumers of news.

**EF: Historical newspaper markets also were highly competitive, with hundreds of U.S. cities having multiple independent daily newspapers. Do competitive media markets offer more, or less, ideological diversity?**

**Shapiro:** We did a series of studies that look at the late 19th century and early 20th century, when it was common for newspapers to have explicit party affiliations. That allows us to get a really sharp lens on what forces seem to be determining those affiliations. The focus for Matt, Mike Sinkinson, and me in one paper was on what generates diversity: When does a city get a Republican-affiliated paper *and* a Democrat-affiliated paper?

That might be a desirable thing to have because it means some views are more likely to come out in discourse. Another reason it might be important is that newspapers are supposed to serve as a watchdog on the state. If there's a Republican in office, the Democratic papers will be vigorously pursuing all of the bad stuff that the politicians do, and vice versa. You want to have both to make sure that, whoever is in charge, somebody would like to break a story about some bad stuff they did and that would give them an incentive to do less bad stuff.

Market forces play a big role in generating this kind of diversity. The likelihood that a newspaper that's entering a market will affiliate with the Republican Party is far greater if the preceding newspaper that entered was a Democrat paper than if the preceding paper was a Republican paper. Why? To an economist, that just looks like product differentiation: That's where the remaining market is going to be. And indeed, what we find is that if you look at data on circulation, when a new Republican paper enters, it tends to crowd out the circulation of other Republican papers more than Democratic papers. It turns out that there are similar things going on in the advertising side of the market: In order to capture the most advertising dollars, it's attractive for newspapers to be different from other newspapers in the market.

**EF: You've found limited evidence that elected officials try to influence media markets — except, it seems, when the stakes are especially high, as in the**

## Jesse Shapiro

### ► Present Positions

George S. and Nancy B. Parker Professor of Economics, Brown University; Research Associate, National Bureau of Economic Research

### ► Selected Past Positions

University of Chicago Booth School of Business (2007-2014); Becker Fellow, Becker Center on Chicago Price Theory, University of Chicago (2005-2007)

### ► Education

Ph.D. (2005), Harvard University; A.B. and A.M. (2001), Harvard University

### ► Selected Publications

"Measuring the Sensitivity of Parameter Estimates to Estimation Moments," *Quarterly Journal of Economics*, 2017 (with Isaiah Andrews and Matthew Gentzkow); "Special Interests and the Media: Theory and an Application to Climate Change," *Journal of Public Economics*, 2016; "Do Pharmacists Buy Bayer? Informed Shoppers and the Brand Premium," *Quarterly Journal of Economics*, 2015 (with Bart J. Bronnenberg, Jean-Pierre Dubé, and Matthew Gentzkow); "Competition and Ideological Diversity: Historical Evidence from U.S. Newspapers," *American Economic Review*, 2014 (with Matthew Gentzkow and Michael Sinkinson); "Fungibility and Consumer Choice: Evidence from Commodity Price Shocks," *Quarterly Journal of Economics*, 2013 (with Justine Hastings)

**Reconstruction-era South. Can you talk about those findings?**

**Shapiro:** Matt, Nathan Petek, Mike, and I studied the effect of party control of the government on the political affiliations of local newspapers. A lot of the research in this area tries to look at the effect of newspapers on voters or voting, and we've done some of that too. But in that study, we flip it around and ask, "What's the effect of who's in office on the news media?"

The context was the late 19th and early 20th century United States, where there are a lot of anecdotes that suggest state governments were doing things like giving patronage jobs to sympathetic newspaper owners or giving contracts to print government notices. We looked at cases where control of government shifts across parties and then what happens after to the share of newspaper circulation that's Republican and the prices newspapers charge. We found surprisingly little impact, but as you said, the overall picture masks some interesting variation.

We thought, where were the incentives to manipulate the news especially high, and where were the market forces especially weak? Those are the two opposing forces determining how media bias leans. We took a look at the Reconstruction South. A lot has been written by historians about media in that period. When the Republicans took control

of Southern governorships and legislatures immediately after the war ended, it's pretty clear that they thought it was very important to support Republican newspapers, perhaps to establish a more long-term position in the South than they actually did. At the same time, the market conditions for Democratic-leaning newspapers were not very good. A lot of owners had been wiped out in the war, or had lost property, or had property confiscated. The economy of the South was in terrible shape and in upheaval. So there wasn't a very robust market for newspapers, and at the same time there was a strong incentive to influence the press.

What we find is that during the period of Republican control, the Republican share of the press is surprisingly high given the prevailing opinions of white Southerners of the time. But when the Democrats reassert control of the government in the late 1870s, they at the same time reassert control of the media. So the story that comes out of

the data is the Republican governments were propping up Republican-affiliated newspapers, and that once that prop was removed a lot of them collapsed. They exit the market or they shrink very rapidly in their circulation.

I like that episode because it shows that it really is about incentives. While commercial incentives are sometimes the most important ones, in some times and places they're not. You can look around the world and see all kinds of examples where things look more like the commercial forces are taking a backseat to political ones. What this paper shows in part is that this has happened from time to time in U.S. history too, even though it hasn't been the main story.

**EF: Some people have suggested there's a dark side to media competition — in particular, that the wide proliferation of choice made possible by the Internet might allow people to filter out views that challenge them. To what extent do you think the Internet is causing political polarization?**

**Shapiro:** If you look in particular at news and political media on the Internet, the proliferation of choice can be overwhelming. As we've discussed, we know that people gravitate toward news outlets similar to their own positions and possibly end up in echo chambers with everybody only hearing from other people who agree with them. A good contrast would be the nightly news broadcast of the 1950s — everybody sits down at the same time every day and watches one of three broadcasts. Everybody is really hearing the same things, and everybody has a shared sense of what's important as a result of that.

In a 2011 article, Matt and I approached the question of how fragmented the audience is the same way that we would think about how a residentially segregated city is. In the literature on residential segregation, the approach is to ask: For people of a given race in that city, what fraction of the other people living around them — say in the same zip code or the same census tract — shares the same race? If that number's very high, that's a very segregated place. People are living around other people like them, and they're likely not having very much social contact with people not like them.

We thought the same idea can be applied to the news media. Think of an online news outlet, like a blog, as a neighborhood, and let's measure who's in that neighborhood: What fraction of those people would self-identify as conservative? What fraction would self-identify as liberal? And let's calculate how segregated is this universe, how segregated is the Internet. To what extent are people visiting news sites that are only populated by other people like them ideologically?

We found that the extent of segregation on the Internet is surprisingly low. It's certainly true that people gravitate to like-minded sources. So for example, foxnews.com has a more conservative audience than nytimes.com.

But the Internet is not radically different from traditional media. Take the fraction of the audience on a given news site that is conservative and call that the conservativeness of the site. Then take the website visited by the average conservative on the average day — that website is about as conservative as usatoday.com. Now do that same thing for the average liberal, that's about as liberal as cnn.com. If you were to read those two outlets, you wouldn't find that they're radically different.

In fact, we find that isolation is very rare in the data. We have individual-level data on users on the Internet. People who get all of their news from outlets to the left of, say, the *New York Times* are very unusual. Likewise, people who get all of their news from sites to the right of Fox News are extremely rare. Folks that go to a fringe conservative site like rushlimbaugh.com are more likely to go to nytimes.com than readers of Yahoo News. The people who are consuming niche media are probably pretty politically engaged people, and therefore they want to read a lot of things. So in the end, the picture is a lot more muted than what people have feared.

Let me give you a metaphor that's sometimes helpful: If you imagine shaking hands with a random other person who's looking at the same news site as you on any given day and ask what fraction of those handshakes are going to be between people who hold different ideological views, that fraction is actually higher on the Internet than, say, in your zip code. So you're actually going to have more contact with people different from you online than you would in the physical world of residential geography. Which I think says that some of these concerns have not been realized, at least not yet.

**EF: You found a result that seems similarly comforting, if you will, when looking at social media.**

**Shapiro:** Right. In a recent paper with Matt and Levi Boxell, we said, let's look at measures of political polarization that have increased from 1996 on — that's the period in which it's relevant to talk about the Internet — and see if the data seem consistent with the narrative that the increase in political polarization is driven by the Internet and social media.

Our approach to testing that hypothesis is very simple: We just compare trends in polarization for groups of people that have high or low propensities to use the Internet and social media. Our favorite and most important comparison is with respect to age. People who are 75 years and over rarely use social media and don't report getting a lot of political information online. People who are 18 to 25 frequently use social media and report getting a lot of political information online. So if you thought that social media was contributing to the rise in polarization, what you would expect to see in the data is that polarization is rising especially fast for younger Americans — and if anything, the story is the opposite. The rise in polarization is similar between the relatively

old and the relatively young, and if anything, maybe polarization is rising faster among the relatively old. So in that sense the data don't line up with the hypothesis that social media is driving the rise in polarization.

I think the effect of the Internet on polarization remains an open question. We're arguing that it doesn't appear that social media is accounting for the increase in polarization, but we haven't offered a constructive account of what *is* driving it. Until we have a better understanding of that, it's hard to rule anything out.

Still, it's easy to confuse changes in the mode of delivery with changes in the information itself. The fundamental processes of generating, communicating, and packaging news are still very much the same as the ones at work 100 years ago.

The notion that you could have a customized news site that is tuned into my exact ideological predisposition misses something very important about the economics, which is that somebody has to write those articles. If you look at news sites that cater to really fringe views, they're not very good. Even if you were to have those views, the spelling is off, they'll miss important events, they'll go off for a couple of days, because the economics just don't support having a staff of people producing, say, the neo-Nazi perspective on all the news of the day. There are not enough people and not enough ad dollars or subscriber dollars to support that kind of activity. And so if you actually want to know what happened yesterday, no matter what your views are, you kind of have to go to a mainstream general news site.

**EF: What do you think are the most interesting questions the 2016 election raised about the role of media in democracy, both the events leading up to the election and since?**

**Shapiro:** To me, the most interesting questions surround the increasing sense that people in America are divided in the world that they perceive. There's a group of people that see a world of progress, a lot of exciting technology, and things getting better all the time, and a group of people who see a world of stagnation, frustration, and things getting worse.

I think there are probably a lot of factors at work there, and I'm not so quick to attribute it to the media. The prior work I mentioned on polarization is one way to try to understand that. Maybe the media play a role in that, or it may be driven by more fundamental social forces like the actual changes in the economy rather than how they're portrayed in the media. Or it may be people's understanding of how the economy works. I'm probably not alone in thinking that the election and the events leading up to it laid those differences bare in a way that they had not been laid bare before. I think that's a very complicated but very important set of issues for, not just economists, but all social scientists.

**EF: You've also found that congressional speech has become more polarized over time. How severe is the shift, and what are the implications for elections?**

**Shapiro:** The basic idea of that study, which was with Matt Gentzkow and Matt Taddy, is based on the fact that we know that Republicans and Democrats speak very different languages. I mentioned before "death tax" and "estate tax," but there are tons of examples. We also know, from a lot of narrative evidence and some quantitative evidence, that those differences in language are often not an accident. They are the result of a conscious process to use language strategically to influence how people see policy issues.

What we didn't know is: How new is that phenomenon? Political parties have been divisive forever in U.S. history, and there have always been differences in how the parties talked about things, what issues they emphasized, and what audiences they were talking to. But the extent of the difference feels, to many people, like it's new and greater than before, and we haven't before had a good way to measure that.

So what we did is try to figure out, for every session of Congress and every point in time, how easily a neutral observer could tell whether someone is a Republican or Democrat based on how they talk. We took the entire *Congressional Record* and used computer scripts to turn it into quantitative data about the use of phrases. Then we took the counts of phrases by every speaker and every session of Congress back to the 1870s and fed that through a model of speech. The model can tell us, at every point in time, how informative your speech is about your party.

What we find is that in the 1870s, if I give you a minute of random speech from somebody in Congress, you're going to guess his party correctly about 54 percent of the time, only modestly higher than chance. In the late 1980s, you'd be doing a little bit better, but barely. By the 2000s, the number is closer to 75 percent. Something enormous changes between the late 1980s and the 2000s to cause the parties to diverge tremendously in how they're talking — many more phrases like "death tax" and "estate tax."

The timing of the change coincides with the "Contract with America" and the Republican takeover in the 104th Congress in 1994. That was a watershed moment in political marketing. It showed the power of language to frame a set of issues and craft a narrative that could be very powerful in winning elections and changing policy views. In the wake of that, strategies on both sides crystallized around trying to have a very consistent message and use very consistent language to try and influence how voters saw the issues. I think that's what's reflected in the data.

In terms of implications, one speculative possibility is that the fact that Republicans and Democrats are speaking differently to each other might contribute to hostility. It might make it harder for them to find common ground or recognize positions on which they do agree.

That's not something that we show in the study, but that's one not-so-optimistic possibility suggested by it.

It may also be that political marketing makes it harder for voters to get to the fundamentals of an issue. We certainly know that in survey experiments, how you phrase a policy proposal can really influence how voters react to it. And so if politicians are increasingly using language in that way, is the net effect that it's harder for voters to tell what their real policy views or interests are? I don't know, but this evidence suggests that this would be an important thing to try and understand.

**EF: Finally, let's talk about your work on food stamps with Justine Hastings. This is one of the oldest and most venerated public programs among economists. What prompted you to study it?**

**Shapiro:** SNAP, the Supplemental Nutrition Assistance Program — the successor to the food stamps program — gives people an electronic benefit transfer card, which is like a debit card. It's topped up by the government every month, and unlike a regular debit card it can only be spent on groceries, food that you're going to consume at home.

This program has been especially interesting to economists because it is an area where economic theory is somewhat in tension with the stated goals of the program. It's clear from the design of the program that the intention is for people to eat, to purchase food. You can see lots of quotes from policy circles that confirm that.

But economic theory says this program might not be about food, and the reason is the following: Think about a household that is spending, say, \$300 on food every month. They join SNAP and start receiving a \$200 per month benefit that can only be spent on food. The household could increase total food spending to \$500 per month and leave everything else the same. Another option would be to take the \$200 in cash that has been freed up and spend it on something else it needs. These are low-income households that have lots of needs, not just for food.

Economic theory basically predicts that this program is really a cash benefit program, equivalent to if I gave you a regular debit card that could be spent on anything. The food stamps program shows up in introductory economics textbooks as an example of the fungibility of money, used to teach the idea of budget constraints and indifference curves in an interesting policy context.

Lots of people have studied this program and tried to adjudicate between these two views of the program — is somebody on SNAP going to mainly increase their food spending or are they going to mainly increase spending on other things?

**EF: What did you find?**

**Shapiro:** We study the question using data from a retail panel where we can track the spending behavior of almost

half a million households for six or seven years. That affords us, a little bit like the newspaper case, lots of little experiments that we can aggregate up to get a very sharp lens on the effects of the program.

What we find is that, whereas the textbook view of the program is that out of every SNAP dollar, maybe 10 cents is going to food and the rest going to other things, we find it's more like 50 cents, which is a big difference in terms of the overall impact on households' budgets and spending behavior. We've known for a long time that people categorize money. We think that part of what's going on is that when SNAP comes in, the household sort of puts it in the food part of the budget, rather than say, "Well, let me just take some other money out of that part of the budget and spread it around evenly," the way that the economics textbook says that they should.

It's unusual for economics to make a quantitative prediction — economics is usually about qualitative predictions, like "when prices go up, demand falls." This is an interesting case where we have a very important public policy — the second-biggest means-tested program in the United States, recently enrolling almost one out of every five U.S. households, touching millions of lives — where economic theory says something very different from the rhetoric surrounding the program. And our data say it's not just wrong, it's not that it's 12 percent instead of 10 percent; it's 50 percent instead of 10 percent. It's wrong by a lot.

It's important for economists to try to figure out what we're missing in the way we model consumer behavior. Because if economists want to advise in designing these programs, which we do, then we want to have models that accurately reflect how households are going to react to these benefits. The textbook model doesn't seem to be serving us very well in this case.

**EF: What are you working on next?**

**Shapiro:** I'm working on a bunch of things. Matt and I remain very interested in applying text analysis to understand social phenomena. A question that we have talked about relates to the 2016 Orlando nightclub shooting. Right after, there was a massive difference in how Republicans and Democrats talked about the incident. Some people talked about "a mass shooting," and some people talked about "Islamic terrorism." How did those differences happen? How fast did they happen? Where did they originate? Did the political actors start it and then more grassroots folks followed? Or was it the reverse? And with data sources like Twitter feeds, it may be possible to see it unfold in real time.

I think there are lots of aspects of the way we think that are going to be revealed by what we say or what we write, and Matt and I are interested in using text analysis methods to try to open up some of those topics, both in modern and historical contexts.

**EF**