Independence and Quality of Mass Media in the Internet Age

The project examines how the Internet has transformed the way news is produced and disseminated, namely through its influence on traditional media, and its ultimate effect on media independence and content quality. To this end, I tackle two distinct but intertwined questions. First, I study how lower advertising revenues affect newspapers' organization and content quality by exploiting the staggered introduction of advertising platform Craigslist across the US. Second, I examine how media dependence on advertisers influences news bias by testing the relationship between advertising spending by car manufacturers and coverage of car safety recalls in US newspapers. My results will shed light on the deep transformations the media industry is undergoing and their implications for the quality of democracy.

Evaluation of Democracy Research Group

Project team:

Ruben DURANTE

Ruben Durante is ICREA Research Professor at UPF, Affiliated Professor of the Barcelona GSE, research affiliate of the Center for Economic and Policy Research (CEPR), and Associate Editor of the *Journal of the European Economic Association*. He holds a PhD in Economics from Brown University.

Valentino LARCINESE

Professor of Public Policy London School of Economics

Associated researchers

- Brian Knight (Brown)
- Ananya Sen (MIT)
- Graham Beattie (University of Pittsburgh)
- Milena Djourelova (Universitat Pompeu Fabra)

Part 1

Objective: to investigate how increased competition from online platforms, and the subsequent decline in advertising revenues, have affected the organization and quality of traditional media, particularly local newspapers.

Empirical approach: to isolate the effect of the entry of online competitors from other confounding factors, we look at the staggered introduction of *Craigslist* – the world's largest online platform for classified ads – which quickly disrupted the local market for classified ads, that had until then been a lucrative niche for local newspapers. In particular, our empirical strategy exploits differences in the timing of the introduction of Craigslist local websites across US media markets since 1996, which, conditional on Internet penetration and key population demographics, appears to have been largely idiosyncratic (Kroft-Pope, 2014; Seamans-Zhu, 2014). In addition, it also exploits differences across local newspapers in pre-Craigslist reliance on classified ads, a measure of the outlet's vulnerability to the new entrant. Specifically, we will implement a *difference-in-difference* approach, by comparing the evolution of newspapers that relied more and less heavily on classified ads ("treated" vs. "control" groups), between the periods before and after the entry of *Craigslist*. To explore the short- and long-term effects of Craigslist's competition on local papers, we will look at various aspects of their organization: i) advertising rates and circulation, ii) the size and composition of newsroom workforce (based on the number and job titles of newspapers' staff), iii) the distribution of editorial staff across different topical areas (e.g., politics vs. entertainment, and local vs. national news). To assess the impact of online competition on content quality, we will look at the evolution of: i) staff workload (proxied by the number of articles signed by each reporter), ii) accuracy of reporting (measured by the number of corrections of past errors), and iii) originality of content (measured by the presence of text segments copied verbatim from news-wire articles or corporate press releases).

Progress: we are currently advancing on the data collection phase, which, due to the considerable amount and variety of data to be gathered, has been rather tedious and slower than expected.

TABLE 1: ADVERTISING SPENDING AND RECALL-RELATED COVERAGE: EXTENSIVE MARGIN

	(1) P(articles)	(2) P(articles)	(3) P(articles)	(4) P(articles)	(5) P(articles)
Log Ad Spending (previous 2 years)	0.918*** (0.127)	-0.519*** (0.117)	-0.651*** (0.113)	-0.271** (0.106)	-0.217** (0.093)
Log Affected Vehicles			0.296*** (0.019)	0.261*** (0.018)	0.261*** (0.019)
Firm's Share Local Cars			0.331*** (0.120)	0.316*** (0.117)	0.301*** (0.068)
Total Articles			0.027*** (0.005)	0.050*** (0.005)	0.049*** (0.005)
Month FE	No	No	No	Yes	Yes
Newspaper x Firm FE	No	Yes	Yes	Yes	No
Newspaper FE	No	No	No	No	Yes
Firm FE	No	No	No	No	Yes
Observations	131,332	131,332	131,332	131,332	131,332
R-squared	0.007	0.124	0.124	0.129	0.14

Robust standard errors in parentheses clustered by newspaper x firm. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the probability that a newspaper writes one or more articles about the recalls of the vehicles produced by a given manufacturer. To improve legibility, the coefficient of "Log Ad Spending (previous 2 years)", and "Log Affected Vehicles", are scaled up by a factor 10². Controls include the logarithm of the number of potentially affected vehicles, firm share of local car demand and the logarithm of total articles written by the newspaper annually. The unit of observation in our empirical analysis is the newspaper-month-manufacturer.

TABLE 2: MEDIA BIAS AND NEWSPAPER COMPETITION

	(1) P(articles)	(2) P(articles)	(3) Log(articles)	(4) Log(articles)
Log Ad Spending (previous 2 years)	-0.415*** (0.141)	-0.462*** (0.143)	-0.430*** (0.140)	-0.465*** (0.145)
Log Ad Spending x Newspaper Competition (previous 2 years)	0.501*** (0.194)	0.569*** (0.193)	0.536*** (0.201)	0.579*** (0.204)
Controls	Yes	Yes	Yes	Yes
Controls x Newspaper Competition	No	Yes	No	Yes
Month FE	Yes	Yes	Yes	Yes
News x Manufac FE	Yes	Yes	Yes	Yes
Observations	131,332	131,332	131,332	131,332
R-squared	0.168	0.168	0.201	0.202

Robust standard errors in parentheses clustered at the Newspaper-Firm level. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the probability that a newspaper writes one or more articles about the recalls of the vehicles produced by a given manufacturer (columns 1 and 2), and the log (+1) of the number of such articles (columns 3 and 4). To improve legibility, the coefficient of Log2 Year Ad Spending is scaled up by a factor 10² equivalent to scaling down Log2 Year Ad Spending) by 10²). Controls include the logarithm of the number of potentially affected vehicles, firm share of local car demand and the logarithm of total articles written by the newspaper annually. The unit of observation in our empirical analysis is the newspaper-month-manufacturer.

Part 2

Objective:

To study to what extent mass media slant content in favor of advertisers, by studying the relationship between advertising spending by car manufacturers and news coverage of car safety recalls in the United States.

Empirical approach:

To disentangle whether newspapers bias content to meet the demands of advertisers or to cater to the preferences of consumers we look at news coverage of car safety recalls in U.S. newspapers since in this context the preferences of advertisers (for less coverage of discreditable incidents) and those of consumers (for more information about safety risks) should affect content in opposite directions. Our analysis focuses on the 100 largest car safety recalls (in terms of the number of vehicles affected) issued between 2000 and 2014 by the U.S. National Highway Traffic Safety Administration (NTHSA). For each recall, we collect information on the date, make and model(s), the number of vehicles potentially affected, nature of the defect(s) and associated safety risks. Regarding news coverage, for each recall, we collect data on all articles about car recalls published in over 130 U.S. daily newspapers, both national and local, available from the *Newslibrary.com* database. For each article, with the assistance of human analysts and the use of sentiment analysis algorithms, we code key qualitative aspects such as the general tone and mentions to accidents and fatalities. With regard to advertising, we acquire data on monthly advertising spending on all U.S. newspapers, by both national car manufacturers and local car dealers, available from Kantar Media's Ad$Spend. Finally, to assess the impact of consumers' demand on news coverage of recalls, we will collect information on vehicle ownership by make at the media market level, available from the National Household Travel Survey for the years 2001 and 2009. The availability of recall-specific data allows us to estimate the impact of ad spending on news coverage, controlling for advertiser-newspaper fixed effects and brand-specific local demand. The availability of data for a 15-year period allows us to test whether a potential pro-advertiser bias is driven by contemporaneous ad spending or, alternatively, whether building a stable relationship with the media, through years of regular and steady spending, allows advertisers to be more influential when it matters. Crucially, we also investigate whether competition for readers and for advertising revenues by other newspapers and online platforms reduces or exacerbates pro-advertising bias.

Progress and results:

Most of the data collection and data analysis phases have been completed. The results provide strong support for our hypothesis. First, we find that newspapers on which a given manufacturer advertised more over the previous two years are significantly less likely to write about recalls involving vehicles produced by that manufacturer, and conditional on writing about them, publish less articles. This effect is stronger for ore serious recalls, i.e., affecting more vehicles and involving more serious defects. Interestingly, only advertising spending in the 6 to 15 months before a recall appear to influence news content, while spending in the months immediately before or several years prior to a recall has no effect (see Figure 1 below). In addition, the stability of the advertising relationship over the previous two years, and not just the total amount spent, influence news coverage of recalls. We also find that pro-advertiser bias is weaker in markets with more newspapers, which indicates that competition for readers has a disciplining effect. In contrast, we find that newspapers experiencing financial distress due to increased competition for advertising by online platforms are more vulnerable to the pressure of advertisers.

Deliverables:

The results of the analysis have recently been published as NBER Working Paper #23940, and as a CEPR Discussion Paper #12366 (both attached). The paper has recently **been accepted for publication in Management Science.**