A Replication of Angelucci, Charles, and Julia Cagé 2019.

Zidong Yang

December 22 2020

Introduction

In recent decades, the hard-copy newspaper industry has been stroke hard by technology progress. While this direct negative effect on companies produce printed newspapers is obvious, the indirect impact on readers is just as profound. The study on the response of newspaper companies toward the market change was set within a time period where a shock took place on the newspaper industry. The shock refers to a new regulation announced by the French government, so now French can air advertisement on their televisions. There are empirical evidences suggest that this relax on regulation has negative impact on newspaper advertising. This further influences the content and readership of newspapers, since the advertising revenue as an essential component of the total revenue of newspaper producers, can restrain companies' budget on anything when it contracts.

The reactions of newspaper companies are captured under a difference-in-difference setting where the difference of newspaper content before and after the shock of decrease in advertising revenue is measured by the difference in national and local newspapers. The fundamental differences between national and local newspapers is everywhere. Primarily, there is a scale difference. Considering fixed effects on the gap between local and national newspapers, the change in national newspapers is only in terms of relative to local newspaper. In the paper, only two fixed effects are included, one for different newspapers and the other for time trend.

Data

The data used for replicating Angelucci, Charles, and Julia Cagé's result is the a panel dataset they digitized. The panel data consists variables on local and national newspapers in France between 1960 and 1970. There are five areas of dependent variables which will be used to investigate the treatment effect. The price, circulation and revenues of those newspapers were collected from the French Ministry of Information's non-publicly available records in the National archives. This information includes 68 local and all 12 national newspapers in France between 1960 and 1974. However, the data of number of journalists which is a strong indicator for the quality as well as the quantity of journalistic-intensive content of newspapers only include information on 63 out of 68 local and 11 out of 12 national newspapers. As for the variables that take account on the change in advertising revenues, advertising price and quantity data are collected from an annual publication, "Tarif Media". The authors' choice of listed price could potentially over-represent the actual transaction price. Especially under the influence of television advertisement, when newspaper companies adjust their advertising prices, some promotions can easily be overseen. An assumption made here is that the gap between the listed price and transaction price is small and does not diverge over time. It is important to treat the listed price with adjustment in order to increase the accuracy of our predication on changes when fitting the model using this data. Also, to capture treatment effect on newspaper content, the newshole data is used. For each newspapers, the ratio of soft and hard news on their front page is a systematically sample of the third week of March from 1964 to 1972. Here, the authors made an extremely subjective scale to distinguish the soft news and hard news on the front page. The distinguishing scale is technically a spectrum, and the standard varies base on readers' personal preferences. This is something can not be measured, but can be generally described. The readership data used by the paper describes readers' characteristics. The source for this data (CESP) publishes every 5 year, the author chose to digitized only 1957, 1962, 1967, 1968, 1968, 1970, 1972 and 1974.

The dataset gives a general picture of French newspaper industry. The scale of local newspapers is approximately one-third of the national given the daily circulation and revenue data. One important fundamental difference between local newspapers and national for this study is the advertisement content on those two types. "National newspapers rely to a greater extent on advertisements for brands, whose owners may also wish to advertise on television. By contrast, advertisements in local newspapers tend to feature classified ads or promote local establishments." (Angelucci, Charles, and Julia Cagé, 2019) The targeted readers for advertisement on those two kinds of newspapers are different. The national newspapers are under larger impact of the shock in the way that the advertising revenue has more reasons to shrink. There are a lot of missing values in the dataset. This forms a natural concern on the statistic power of the study. Missing values can potential effect the precision of estimated changes due to the shock. However, the authors provided a robustness check to accommodate this concern. The robustness checks will be discussed later in this report.

Model and Results

The effect of the introduction of advertising on television is measure by DiD estimation method. Specifically, use national newspapers as the treatment group, the control group is local newspapers who should be receiving minor impact according to authors' assumption. The DiD estimation replies on the common trend assumption. In context of this study, the assumption can be interpreted as follows. In the absence of the shock created by the introduction of advertising on television, the difference between French national newspapers should be the same before and after 1967. The variables that are taking account for the effect of the shock can be divided into four groups.

Table 1: Advertising Side

	Ad. rev.	Ad rev. div. circ.	Ad price	Ad space
after_national1	-0.23 ***	-0.15 ***	-0.31 ***	0.01
	(0.03)	(0.03)	(0.07)	(0.05)
year	0.05 ***	0.04 ***	0.04 ***	0.02 ***
	(0.00)	(0.00)	(0.00)	(0.00)
N	1052	1048	809	1046
R2	0.99	0.90	0.89	0.72
logLik	345.34	449.52	-277.71	-164.01
AIC	-526.68	-735.05	705.43	478.02
*** p < 0.001; ** p < 0.01; * p < 0.05.				

The group representing the effect on advertising side (Table 1) for the newspapers are advertising revenues, listed advertising price and advertising space. Since the revenue is highly dependent on the circulation of the newspaper, a circulation-wide normalization is used on advertising revenue forming a new variable.

Another group of the dependent variables is called reader side. The reader side group (Table 2) explains the indirect effect of the shock on readers. The replicated results of these two groups is somewhat consistent with the paper. The replicated amount of percent change in advertising price of national newspapers compared to the price of local newspaper is -31% instead of -40% as showed in paper. Also, the effect on advertising space is positive in my replication while the paper gives a negative three percent change. However, this change is

Table 2: Reader Side

	Subscription price	Unit price	Circulation
after_national1	-0.04 *	0.06 **	-0.06 **
	(0.02)	(0.02)	(0.02)
year	0.05 ***	0.05 ***	0.01 ***
	(0.00)	(0.00)	(0.00)
N	1044	1063	1070
R2	0.88	0.87	0.99
logLik	882.14	907.28	759.57
AIC	-1600.28	-1650.57	-1355.15
*** p < 0.001; ** p < 0.01; * p < 0.05.			

still statistically insignificant. My model gives a smaller estimation for the effect on listed advertising price as well as advertising space. (with a overall smaller standard error close to zero).

The quality of one newspaper are measured by one's number of journalists and newshole. There are a lot of missing data on the payroll variable, which weakens the power of estimation. The replication is fairly consistent with the paper. The change in share of hard news on front page is estimated 4% bigger than the paper. This could be because the low number of observations. Specifically, the data of the hard news ratio on front page only contains 37 local newspapers and 10 national newspapers from 1964 to 1972. A lower observation weakens the prediction model.

Table 4 and 5 is the replication result for the effect of decreased advertising revenue on readership composition. Due to lack of local newspaper dataset in readership, the results are only suggestive here. Notice the shift in the readership composition, which can be explained by the shift in newspaper content from hardcore news to soft. A smaller ratio of hard news on the front page could be less attractive to people with higher education. Or the newspapers "chose" to shift their focus to readers with lower education to lower their cost on content producing. This result provides consistency to the picture drawn by previous results.

Table 6 take in account of heterogeneous effect in which the extend of reliance on the advertising revenue are different before the introduction of advertising on television. To show newspapers with higher reliance on advertising revenue before suffers stronger effect, an OLS estimation was run on different dependent variables in newspaper industry for both the low reliance kind and high. The benchmark for high reliance is the median for reliance on advertising revenue. After we drop the low reliance and local newspapers, there are only 91 observations left in the dataset including the ones with missing values. The replicated result is not very statistically significant, so the interpretation is unnecessary. The replicated results are off. The paper did not mention the procedure to reproduce the data used for investigating the heterogeneous effect. The results showed in the paper, however, does confirm the a higher impact on newspapers with high reliance on advertising revenue. Except for one variable, the subscription price which shows a bigger drop in newspapers with low reliance. The author did not mention why. A possible explanation for this matter form me is that newspapers who did not rely much on the advertising revenue might rely on something else more than those who rely more on advertising. What if they rely more on their revenue form subscriptions? If so, in order to attract more subscribers in the difficult time, they have to give bigger discount on the subscription price.

To check the validity of DiD estimation, the model interacts the year fixed effect with national newspaper indicator variable. The coefficients replicated are all very small, which is just as expected. However, for

Table 3: Reader Side

	Share of sub	Revenue from sales
after_national1	0.19 ***	-0.06 *
	(0.03)	(0.03)
year	-0.01 ***	0.05 ***
	(0.00)	(0.00)
N	1072	1046
R2	0.97	0.99
logLik	321.91	451.11
AIC	-477.81	-738.22
*** p $<$ 0.001; ** p $<$ 0.01; * p $<$ 0.05.		

Panel G and H, the coefficients show something interesting. The coefficient of the interact term in Panel G suggests that employee share in readership decreased as suggested in the Table 5. The coefficient of the interact term in Panel H suggests an increase in readers with lower educations as suggested in the Table 5. The negative 0.06 in Panel B can easily be explained by newspapers' adjustment under the impact of television advertising. The results are consistent the paper.

Discussion

The paper has an out-standing robustness check on the results. One interesting check is Weighting News-papers by Their Circulation. Is circulation the only worthy factor when weighting the newspapers? The circulation of newspapers itself dependents on many other factors. While the individual characteristic of each newspaper is capture in terms of assigning each newspaper with an ID, there are different things happening in different industries. Therefore newspapers for a specific industry, for example politics, should be accommodate with specific time trend. An Industry-Specific Time Trend robustness check was also provided by the author. This is necessary especially in the context of France as in 1960s, the politics is very active in France. The printed press were free, but the television and broadcasts were under government control (Kuhn, 2006). The preference between newspaper and television might shift when the government is under an unstable condition. The choice of accessing information depends on not only the time, but what was happening at the time. Of course we cannot add fixed time effect whenever something happened that caused a shift in people's preferences. The best that was done is by adding the annual adjustment to the model. However, the unstable politics could potentially offset the changes in newspaper industry due to the shock of 1967.

The mean unit price of national newspaper is 3.6 euros which is higher than local newspapers. The decreased subscription price of national newspapers make it more accessible to people who were unable to afford. This group of people who probably haven't read national newspapers before are now able to afford. If they didn't even know what was the content of national newspapers before, how would they subscribe due to the adjustment in contents? The factor that triggered them to subscribe is just the affordable price to them. After they joined the readership of national newspapers, they made up the increased portion on lower educational level readers because they have relatively lower income. Therefore we cannot interpret the shift in readership composition is purely because of the "lower quality" of the national newspapers. There is an asymmetric information issue to the new subscribers.

Table 4: Quality

	Number of journalists	Average payroll	Number of pages
after_national1	-0.20 ***	0.06	-0.02
	(0.03)	(0.05)	(0.02)
year	0.04 ***	0.03 ***	0.03 ***
	(0.00)	(0.00)	(0.00)
pqr_trend			
N	1046	723	1046
R2	0.98	0.56	0.93
logLik	381.07	12.15	816.27
AIC	-610.14	81.71	-1482.54
*** p < 0.001; ** p < 0.01; * p < 0.05.			

Also, the decrease in the portion of artisans and shopkeepers have other alternative explanations. Considering the relatively free working environment of shopkeepers, it is likely that they have access to news during work already. When local advertising companies aired on television, some useful pecieces of information were able to be received more efficient for someone who already have access of television or radio. For example, some convenience stores have their radio on in the background, restaurants sometimes have television installed to attract more customers as well. When the advertising on television was introduced,they might choose to save the money on subscribing newspapers. When this situation is considered, the percent decrease in artisans and shopkeepers share of the national readership is actually over-estimated. One further improvement can be taken on precision and accuracy of estimation is solving for the missing data. There are too many missing values especially for the dependent variable, share of subscribers. Even the robustness check on Balanced Sample provided by the author failed on this variable. The data collecting process for this study was challenging. Even with much more advanced technology, if one want to mine data as ancient as the time period investigated in the paper, digitization is still an indispensable step.

There are some other independent variables can be added to the OLS model. Aside of time and newspaper fixed effects, geographic effect can be added and used to explain the unit sale and number of subscribers. Moreover, the investigation on different national and local newspapers for different industries can also add some accuracy to the model. For example, we can categorize financial newspapers as one group and run OLS estimation on them with more precised fixed effect variables, and then political newspapers as another and so on.

References

Angelucci, Charles, and Julia Cagé. 2019. "Newspapers in Times of Low Advertising Revenues." American Economic Journal: Microeconomics, 11 (3): 319-64.DOI: 10.1257/mic.20170306 David Hugh-Jones (2020). huxtable: Easily Create and Style Tables for LaTeX, HTML and Other Formats. R package version 5.1.1. https://CRAN.R-project.org/package=huxtable H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016. Hadley Wickham and Dana Seidel (2020). scales: Scale Functions for

Table 5: Quality

	Number of journalists	Average payroll	Number of pages
after_national1	-0.20 ***	0.06	-0.02
	(0.03)	(0.05)	(0.02)
year	0.04 ***	0.03 ***	0.03 ***
	(0.00)	(0.00)	(0.00)
pqr_trend			
N	1046	723	1046
R2	0.98	0.56	0.93
logLik	381.07	12.15	816.27
AIC	-610.14	81.71	-1482.54
*** p < 0.001; ** p < 0.01; * p < 0.05.			

Visualization. R package version 1.1.1. https://CRAN.R-project.org/package=scales Hadley Wickham and Evan Miller (2020). haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files. R package version 2.3.1. https://CRAN.R-project.org/package=haven Kirill Müller (2020). here: A Simpler Way to Find Your Files. R package version 1.0.1. https://CRAN.R-project.org/package=here Raymond Kuhn (2006). The Media in France. Routledge. pp. 70–180. ISBN 9781134980536. Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686 Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686

Table 6: Table 4 Readership: Eduction

	Nodiploma	Primaryeducation	Secondaryeducatiion	Professionaleducaition
after_national1	0.52	1.78 **	-2.26 ***	0.14
	(0.71)	(0.69)	(0.64)	(0.58)
year	-0.12 *	-1.29 ***	0.58 ***	0.82 ***
	(0.05)	(0.05)	(0.04)	(0.04)
N	413	413	413	413
R2	0.36	0.96	0.79	0.94
logLik	-1053.11	-1042.43	-1015.14	-975.28
AIC	2188.22	2166.86	2112.29	2032.55
*** p < 0.001; ** p < 0.01; * p < 0.05.				

Table 7: Quality

	Farmers	Artisansand shopkeepers	Senior executives
after_national1	2.50 ***	-0.11	-1.01 **
	(0.49)	(0.33)	(0.37)
year	-0.54 ***	-0.14 ***	0.08 **
	(0.03)	(0.02)	(0.02)
N	413	413	413
R2	0.94	0.71	0.96
logLik	-899.33	-733.41	-784.40
AIC	1880.67	1548.82	1650.80
*** p < 0.001; ** p < 0.01; * p < 0.05.			

Table 8: Quality

	Employees	Labourers	Inactive
after_national1	-8.73 ***	4.83 ***	0.32
	(0.61)	(0.85)	(0.48)
year	-0.39 ***	-0.32 ***	0.87 ***
	(0.04)	(0.06)	(0.03)
N	413	413	413
R2	0.77	0.90	0.86
logLik	-992.58	-1132.09	-890.99
AIC	2067.15	2346.19	1863.98
*** p < 0.001; ** p < 0.01; * p < 0.05.			

Table 9: Heterogenous Effects: Reliance on Advertising of National Daily Newspapers before the Shock

	Ad revenue		(Listed) ad price		Subscription price	
	Low	**High**	**Low**	**High**	**Low**	**High**
after_national1	-0.01	-0.16	-0.02	0.05 *	0.16 ***	0.20 ***
	(0.12)	(0.10)	(0.08)	(0.02)	(0.03)	(0.04)
year	0.04 *	0.02	0.01	-0.03 ***	0.01 **	0.02 **
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)
N	77	84	72	49	79	69
R2	0.97	0.96	0.95	0.99	0.82	0.90
logLik	-5.55	5.98	29.09	84.10	93.58	68.83
AIC	29.09	6.04	-40.18	-154.20	-169.17	-121.65
*** p < 0.001; ** p < 0.01; * p < 0.05.						

Table 10: Heterogenous Effects: Reliance on Advertising of National Daily Newspapers before the Shock

	Share of subscribers		Number of journalists		Newshole	
	Low	**High**	**Low**	**High**	**Low**	**High**
after_national1	-0.01	-0.06	0.01	0.00	0.01	0.00
	(0.13)	(0.10)	(0.05)	(0.13)	(0.08)	(0.06)
year	0.02	0.03 **	0.02 *	0.01	0.03 **	0.01
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
N	79	69	87	71	84	54
R2	0.92	0.98	0.91	0.91	0.76	0.50
logLik	-15.05	14.99	63.16	-3.71	25.34	44.05
AIC	48.09	-13.99	-108.31	23.41	-32.69	-74.10
*** p < 0.001; ** p < 0.01; * p < 0.05.						

Table 11: Controlling For Parallel Trend

	Panel A. Advertising revenues	**Panel B. Advertising price**
national1	59.08 ***	106.82 ***
	(7.55)	(19.73)
year	0.05 ***	0.05 ***
	(0.00)	(0.00)
national1:year	-0.03 ***	-0.05 ***
	(0.00)	(0.01)
N	1052	809
R2	0.99	0.89
logLik	349.83	-271.74
AIC	-535.66	693.47
*** p < 0.001; ** p < 0.01; * p < 0.05.		

Table 12: Controlling For Parallel Trend

	Panel C. Subscription price	Panel D. Circulation
national1	27.15 ***	18.92 ***
	(4.64)	(5.13)
year	0.05 ***	0.01 ***
	(0.00)	(0.00)
national1:year	-0.01 ***	-0.01 ***
	(0.00)	(0.00)
N	1044	1070
R2	0.88	0.99
logLik	898.40	761.50
AIC	-1632.80	-1358.99
*** p < 0.001; ** p < 0.01; * p < 0.05.		

Table 13: Controlling For Parallel Trend

	Panel E. Number of journalists	Panel F. Newshole
national1	58.22 ***	12.29 *
	(6.99)	(5.41)
year	0.04 ***	0.03 ***
	(0.00)	(0.00)
national1:year	-0.03 ***	-0.01 *
	(0.00)	(0.00)
N	1046	1046
R2	0.98	0.90
logLik	391.89	742.04
AIC	-631.77	-1334.08
*** p < 0.001; ** p < 0.01; * p < 0.05.		

Table 14: Controlling For Parallel Trend

	Panel G.Readership:Percent employees	Panel H.Readership:Percent farmersand lab
l1	1705.01 ***	-145
	(166.85)	(1
	-0.40 ***	-
	(0.05)	
l1:year	-0.86 ***	
	(0.08)	
	413	4
	0.73	
	-1032.85	-11
	2147.69	22

< 0.001; ** p < 0.01; * p < 0.05.

11