How Increased Chinese Exports Drive Media Slant? Evidence from U.S. Local Newspaper over 1998-2017

Linghui Wu

May 26, 2020

Abstract

Does the recent surge in Chinese imports affect the media slant against China in the United States? Using a dataset of 157 U.S. local newspapers from 1998 to 2017, we construct a new measure of media slant with the LSTM sentiment algorithm. The paper shows that newspapers whose circulation states face greater exposure to Chinese import shocks report more negative news about China. The source of negative descriptions more stems from non-trade-related topics rather than trade-related articles. Further, the increase of female and Asian population shares restrain the rise of negative trade-related coverage. The results are robust with an alternative measurement of media slant with TextBlob in Python.

1 Methods and Results

1.1 Data

The paper focuses on U.S. daily newspapers with state-level circulation data which are available from 1998 to 2017. Newspaper county-level circulation data are retrieved from *Alliance of Audited Media*, the largest North American non-profit organization providing verified media information. Newspaper contents are obtained from *Newslibrary* by dynamically web scraping the abstracts. Due to copyright, the database only allows viewing the first 500 characters (around 90 words) of each article. Those words usually form the leading paragraph where writers commonly express their opinions. Overall, the data collected contain 114,788 pieces of articles from 157 local newspapers, covering 52 U.S. states (see Table 3 in Appendix A for details).

4-digit Harmonized System level international trade data are from *U.N. Comtrade Database* and data on state-level industry structure and demographics are from *U.S. Census Bureau*.

1.2 Key Variables

1.2.1 Media Slant

The study is concerned with media slant, which is caused by selective coverage of negative topics and biased descriptions through prejudiced expressions that would project a negative image of China. The definition is commonly discussed in literatures (Groseclose and Milyo, 2005; Gentzkow and Shapiro, 2010).

To measure media slant, previous studies tend to construct a list of negative keywords and then use it to identify whether a report is negative or not (Larcinese et al., 2011; Puglisi and Snyder Jr, 2011; Ramirez and Rong, 2012; Lu et al., 2018). However, keyword detection is unreliable because of not only the subjectivity in establishing the negative-word dictionary but also the ambiguity - not all reports that mention the "negative keywords" stigmatize China. Therefore, this paper adopts the long short-term memory (LSTM) algorithm for text sentiment analysis. LSTM is is a chain-structured neural network architecture extended from the recurrent neural network (RNN). It is composed of input gates, output gates and forget gates to filter information passed down from upper layers, which contributes to its outperformance in longer input sequences.

The training set comes from a large movie review dataset provided by Al Lab at Stanford University (Maas et al., 2011). The binary sentiment classification dataset consists of 25,000 pieces of data, and labels positive reviews as 1 and negative ones as 0. In the pre-processing, we employ nltk package to tokenize and vectorize the text so that human languages can be comprehended by the algorithm. A list of English stop words including "a", "an", "to" is removed. Then we import keras library in Python to build the LSTM neural network, which is implemented on the basis of the Tensorflow framework. The main parameters of the model after tuning are shown in Table 4 in Appendix B. The output is a real number in [0, 1] and by convention, 0.5 is chosen as the classification threshold. The trained sentiment classification model based on deep learning algorithm reaches around 82% accuracy in the test set.

We resort to *Newslibrary* database for China-related articles over 1998 to 2017 by locating the reports with "China" or "Chinese" in headlines, and obtain the total number of news articles about China by newspaper i in year t, denoted as $China_{i,t}$. Then, we apply the trained LSTM sentiment classification model to the selected articles and attain the number of articles with negative reporting about China, denoted as $Negative_{i,t}$. As a result, the ratio of negative articles about China to the total number of China-related articles is the measurement of media slant against China.

$$NegRatio_{i,t} = \frac{Negative_{i,t}}{China_{i,t}} \tag{1}$$

As shown in Table 5, the average change in media slant from 1998 to 2017 ($\Delta NegRatio_i$) is 0.306 with a standard deviation of 0.289.

For sub-sample analysis, we hope to investigate what topics the newspapers emphasize when depicting a negative image of China after suffering the imported Chinese trade shocks. To this end, we further divide the news reports into trade-related and non-trade-related ones, and construct two additional measures for media slant against China; that is, one for trade-related, $NegRatio_{i,t}^{trade} = \frac{Negativei,t^{trade}}{China_{i,t}}$, and the other for non-trade-related, $NegRatio_{i,t}^{non-trade} = \frac{Negative_{i,t}^{non-trade}}{China_{i,r}}$.

The aforementioned LSTM measurement of media slant has potential shortcomings since movie reviews and newspaper reports belong to distinct genres of writings and the different distribution of high-frequency words may lead to bias in predicting. Accordingly, in robustness checks, we provide an alternative sentiment classification via the TextBlob library in Python, which is more compatible with various kinds of English text. Section 2.3 provides details on the measurements and results.

1.2.2 Imported Trade Shocks from China at the Newspaper Level

Because the dependent variable $NegRatio_{i,t}$ concerns the reporting behavior of newspapers, it is required to measure the independent variable, import exposure to China, and a host of control variables at the newspaper level.

To create the regressor of interest, we follow Autor et al. (2013) in two steps. First, we construct state-level changes in Chinese imports using industry-level data from *U.N. Comtrade Database*[1] and state-level employment structure data from *U.S. Census Bureau*. Second, we use the newspaper circulation data across states as weights and sum the changes in Chinese imports calculated in the first step for the newspaper-level measurement of Chinese imports. Mathematically, the measurement is given by:

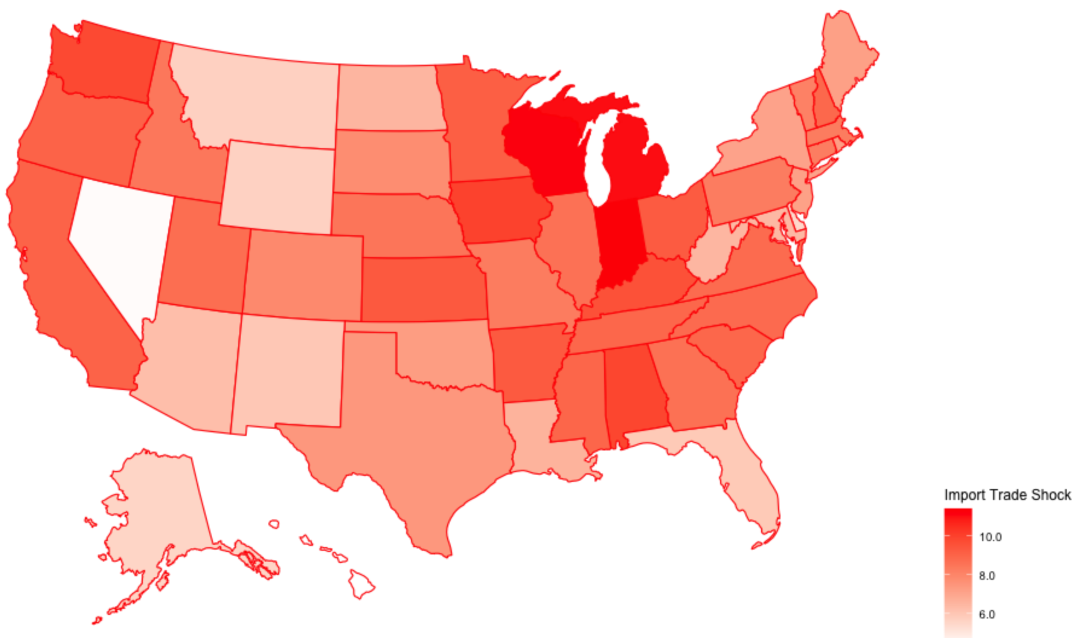
$$\Delta Import_m^{China} = \sum_s \frac{Circulation_{i,s}^{1998}}{Circulation_i^{1998}} \sum_s \frac{L_{s,j}^{1998}}{L_s^{1998}} \frac{\Delta M_j^{China}}{L_j^{1998}} \tag{2}$$

where ΔM_j^{China} is the change in U.S. imports from China between 1998 to 2017 in industry j; $L_{s,j}$ is the employment in industry j in state s in 1998; L_s is the employment in state s in 1998; L_j is the employment in industry j in the United States in 1998; $Circulation_{i,s}^{1998}$ is the weekly circulation of newspaper i in state s in 1998; and $Circulation_i^{1998}$ is the total circulation of newspaper i in 1998. Note that if it is not due to the constraints of data availability, it produces more accurate

[1]We extract 4-digit HS96 trade data from *U.N. Comtrade* and then convert the trade data to 2-digit NAICS (North American Industry Classification System).

results by replacing calibration variables with time-varying values. Figure 1 shows the change in Chinese import competition calculated by the above Autor et al. (2013) method from 1998 to 2017 across U.S. states, with darker color indicating greater shocks from Chinese imports.

Figure 1: U.S. Exposure to Chinese Imports across States, 1998-2017



U.S. exposure to Chinese imports across states from 1998 to 2017 calculated according to Autor et al. (2013). Darker color indicates greater increase in Chinese import competition from 1998 to 2017.

2 Empirical Findings

2.1 Specifications

To investigate the effect of exposure to Chinese imports on media slant, we follow the strategy proposed in Autor et al. (2013), which probes variations in state exposure to Chinese imports. The estimation specification is as follows:

$$\Delta NegRatio_{i,t} = \alpha + \beta \Delta Imports_{i,t}^{China} + \lambda \boldsymbol{X}_{i,1998} + \Delta\epsilon_{i,t} \tag{3}$$

where $\Delta NegRatio_{i,t} \equiv NegRatio_{i,t} - NegRatio_{i,1998}$ captures the change in media slant against China by newspaper i from 1998 to year t; similarly, $\Delta Imports_i^{China}$ measures the change to Chinese imports in the circulation states of newspaper i from 1998 to year t; $\Delta\epsilon_{i,t}$ is the error term. To mitigate the possible relationship between the independent variable $\Delta NegRatio_{i,t}$ and states' industrial and newspaper circulation structures, we compute all the weights in the early periods for which we have access to data, such as newspapers' circulation distribution across states in 1998 and employment statistics by state in the 1990s. We also include a vector of circulation-weighted shares of the readership attributes $\boldsymbol{X}_{i,1998}$: female population, Asian population, population with a bachelor's degree and median income level.

This difference operation in Equation 3 helps eliminate newspaper fixed effect; or in other words, the analysis controls for all time-invariant differences across newspapers. Meanwhile, the identification in Equation 3 is from the cross-newspaper variations in the same time period, which helps to control for the time effects that are common to all newspapers such as the possible improvement or deterioration of the social, culture, or political environment in China. However, the potential estimation biases of $\Delta Imports_{i,t}^{China}$ could stem from the endogenous change in Chinese imports from 1998 to year t, ΔM_j^{China}, and the nonrandom distribution of industrial structure and newspaper state-level circulation, $\frac{Circulation_{i,s}^{1998}}{Circulation_i^{1998}}$ and $\frac{L_{s,j}^{1998}}{L_s^{1998}}$.

2.2 Baseline Results

The estimation results are reported in column (1) and (2) in Table 1, without and with additional controls. The regressions suggest positive and statistically significant coefficients of change in Chinese imports, suggesting the exposure to Chinese imports causes newspapers in the United States to report more negative news about China.

Table 1: Baseline Results

	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta NegRatio$	All Sample	All Sample	Trade	Trade	Non-trade	Non-trade
$\Delta Imports$	0.0283***	0.0281***	0.00999***	0.00974***	0.0286***	0.0284***
	(17.80)	(17.28)	(6.13)	(5.80)	(17.89)	(17.32)
Female		0.131		-0.218*		0.130
		(0.89)		(2.24)		(0.88)
Asian		-0.00582		-0.308*		-0.0709
		(-0.05)		(-2.50)		(-0.59)
Bachelor		-0.0110		0.389		0.0511
		(-0.05)		(1.83)		(0.25)
Income		-0.00616		-0.0293		-0.00818
		(-0.30)		(-1.32)		(-0.39)
Constant	0.157***	0.119*	0.104***	0.00173	0.161***	0.116*
	(16.38)	(2.10)	(11.60)	(0.23)	(16.68)	(2.03)
N	3140	3140	3140	3140	3140	3140

Note: t-statistics in parentheses. The dependent variables in column (1)-(2), (3)-(4) and (5)-(6) are the change in the percentage of newspaper's negative reports in total China-related reports, the change in the percentage of negative reportings about trade-related news in total China-related articles and the change in the percentage of negative reportings about non-trade-related news in total China-related articles from 1998 to 2017. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

We have shown that exposure to Chinese imports increased negative reporting on China by U.S. local newspapers. It would be interesting to know the source of such negative reports about China. Hence, we divide all China-relevant negative articles into two parts (trade-related and non-trade-related), and construct two sub-components of media slant. One subgroup is the percentage of negative trade-related reports in all China-related articles ($NegRatio_{i,t}^{trade} = \frac{Negative_{i,t}^{trade}}{China_{i,t}}$) and the other is the percentage of the negative non-trade-related reports in all China-related articles ($NegRatio_{i,t}^{non-trade} = \frac{Negative_{i,t}^{non-trade}}{China_{i,t}}$). Next, we investigate the impact of Chinese import shocks on each of the two subgroups of media slant.

Table 1 shows that the change in the Chinese imports exposure significantly impact both the change in the percentage of negative trade-related reportings on China (column (3) and (4)) and that of negative non-trade-related reportings on China (column (5) and (6)). However, the effect of trade-related news (0.00974) is smaller than that of non-trade-related news (0.0284). These results indicate that most of the increased negative newspaper articles about China rise from non-trade topics such as human rights, the political regimes, which align with the previous studies (Larcinese et al., 2011; Ramirez and Rong, 2012; Lu et al., 2018) that partisan bias in newspaper coverage is less biased for trade issues than other economic issues.

One possible explanation is that, compared with trade-related reports, it is more direct and easier for newspapers to express negative attitudes against China on ideological topics. Another reasoning might boil down to the fact that newspaper's coverage of trade issues is driven by special interest groups. Groups such as labor unions and environmentalists are under-represented in newspaper coverages and on the contrary, newspaper articles largely depended on interviews with business representatives, who were generally pro-trade (Baker, 1994).

Apart from the independent variable, female and Asian population are negative and statistically significant, implying that the increase of female and Asian population in the newspaper-circulated areas decreases the ratio of negative trade-related reportings to the total China-related articles. The coefficients are reasonable in that White males disproportionately occupy the sectors that undergo the Chinese import shocks the most (McManus and Schaur, 2016; Pierce and Schott, 2020).

2.3 Robustness Checks

The aforementioned LSTM media slant may have potential measurement errors caused by the distinctions between movie reviews and newspaper articles. For robustness checks, we employ the TextBlob in Python to perform sentiment classification. The package has wide application because of the simple dictionary-based algorithm. As shown in Table 5, for all 157 newspapers, the average change in negativity from 1998 to 2017 is 0.482 with a standard deviation of 0.381.

Table 2: Robustness Check

	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta NegRatio$	All Sample	All Sample	Trade	Trade	Non-trade	Non-trade
$\Delta Imports$	0.0457***	0.0466***	0.0178***	0.0173***	0.0464***	0.0474***
	(22.18)	(22.21)	(7.62)	(7.19)	(22.42)	(22.48)
Female		0.0872		-0.390**		0.0676
		(0.38)		(2.90)		(0.29)
Asian		-0.408*		-0.0955		-0.327*
		(2.44)		(-0.56)		(1.97)
Bachelor		0.203		0.245		0.356
		(0.80)		(0.87)		(1.41)
Income		-0.0515*		-0.0245		-0.0594*
		(-1.99)		(-0.83)		(-2.30)
Constant	0.242***	0.304**	0.177***	0.00442	0.244***	0.302**
	(18.32)	(3.09)	(13.38)	(0.43)	(18.50)	(3.07)
N	3140	3140	3140	3140	3140	3140

Note: t-statistics in parentheses. The dependent variables in column (1)-(2), (3)-(4) and (5)-(6) are the change in the percentage of newspaper's negative reports in total China-related reports, the change in the percentage of negative reportings about trade-related news in total China-related articles and the change in the percentage of negative reportings about non-trade-related news in total China-related articles from 1998 to 2017. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 2 reports the results for the Blob-measured media slant. The coefficients in column (1) and (2) indicate a positive and statistically significant impact of Chinese imports on media slant. Robustness checks are also applied to sub-sample analysis and we can derive similar conclusions from column (3) to column (6): the increase in Chinese import exposure significantly rise the amount of negative trade-related and negative non-trade-related reportings to the total negative China-relevant articles; the impact of trade deficit shocks are more substantial on the non-trade-related coverage (0.0474) than the trade-related coverage (0.0173); newspapers whose circulation states have higher female population are less likely to cover negative trade-related news.

References

Autor, David H, David Dorn, and Gordon H Hanson, "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *The American Economic Review*, 2013, pp. 2121–2168.

Baker, Dean, "Trade Reporting's Information Deficit," *EXTRA/(November*, 1994.

Gentzkow, Matthew and Jesse M Shapiro, "What drives media slant? Evidence from US daily newspapers," *Econometrica*, 2010, *78* (1), 35–71.

Groseclose, Tim and Jeffrey Milyo, "A measure of media bias," *The Quarterly Journal of Economics*, 2005, *120* (4), 1191–1237.

Larcinese, Valentino, Riccardo Puglisi, and James M Snyder Jr, "Partisan bias in economic news: Evidence on the agenda-setting behavior of US newspapers," *Journal of public Economics*, 2011, *95* (9-10), 1178–1189.

Lu, Yi, Xiang Shao, and Zhigang Tao, "Exposure to Chinese imports and media slant: Evidence from 147 US local newspapers over 1998–2012," *Journal of International Economics*, 2018, *114*, 316–330.

Maas, Andrew L., Raymond E. Daly, Peter T. Pham, Dan Huang, Andrew Y. Ng, and Christopher Potts, "Learning Word Vectors for Sentiment Analysis," in "Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies" Association for Computational Linguistics Portland, Oregon, USA June 2011, pp. 142–150.

McManus, T Clay and Georg Schaur, "The effects of import competition on worker health," *Journal of International Economics*, 2016, *102*, 160–172.

Pierce, Justin R and Peter K Schott, "Trade liberalization and mortality: evidence from US counties," *American Economic Review: Insights*, 2020, *2* (1), 47–64.

Puglisi, Riccardo and James M Snyder Jr, "Newspaper coverage of political scandals," *The Journal of Politics*, 2011, *73* (3), 931–950.

Ramirez, Carlos D and Rong Rong, "China Bashing: Does Trade Drive the "Bad" News about China in the USA?," *Review of International Economics*, 2012, *20* (2), 350–363.

A Newspaper List

Table 3: List of 157 U.S. Local Newspapers

Newspaper Name	Newspaper Name	Newspaper Name
Adelante Valle	Journal Inquirer	Quad-City Times
Aegis	Journal Star	Rapid City Journal
Albany Democrat-Herald	Kalamazoo Gazette	Record-Courier
Albuquerque Journal	Kane County Chronicle	Republican-American
Amarillo Globe-News	Kearney Hub	Richmond Times-Dispatch
American Press	Kent County Daily Times	San Antonio Express-News
Anacortes American	Kerrville Daily Times	San Francisco Chronicle
Anniston Star	Knoxville News Sentinel	Santa Fe New Mexican
Athens Banner-Herald	Laredo Morning Times	Saratogian
Bainbridge Island Review	Las Cruces Sun-News	Savannah Morning News
Beaver County Times	Las Vegas Review-Journal	Skagit Valley Herald
Carlsbad Current-Argus	Lincoln Journal Star	Soundoff!
Carroll County Times	Lodi News-Sentinel	South Jersey Times
Chicago Sun-Times	Longview News-Journal	St. Joseph News-Press
Citrus County Chronicle	Lubbock Avalanche-Journal	St. Louis American
Columbus Parent Magazine	Martinsville Bulletin	St. Louis Post-Dispatch
Connecticut Post	Mechanicsville Local	St. Paul Pioneer Press
Corpus Christi Caller-Times	Mercury	Standard Times
Coventry Courier	MetroWest Daily News	Standard-Examiner
Culpeper Star-Exponent	Midland Daily News	Stanwood Camano News
Cuyahoga Falls News-Press	Milford Daily News	Star-Herald
Daily Chronicle	Milwaukee Journal Sentinel	Stow Sentry
Daily Citizen	Missoulian	Sun Advocate
Daily Comet	Morning News	Sun Journal
Daily Inter Lake	Mountain Democrat	Tallmadge Express
Daily Item	Mountain Eagle	Tampa Bay Times
Daily Journal	Napa Valley Register	Taunton Daily Gazette
Daily Local News	Naples Daily News	Telegraph
Daily News	Narragansett Times	Texarkana Gazette
Daily Press	New Braunfels Herald-Zeitung	The Palm Beach Post
Daily Republic	New Haven Register	The Philadelphia Inquirer
Daily Sentinel	New York Post	The Record
Davis Enterprise	News Journal	Times Herald
Dayton Daily News	News Leader	Times Herald-Record
Delaware County Daily Times	Niagara Gazette	Times Union
Delaware State News	Nordonia News-Leader	Times-News
East Greenwich Pendulum	Northeast Mississippi Daily Journal	Twinsburg Bulletin
El Paso Times	Northern Virginia Daily	Tyler Morning Telegraph
Gleaner	Northwest Florida Daily News	Valley Morning Star

Goochland Gazette	Northwest Herald	Ventura County Star
Hartford Courant	Observer-Dispatch	Voices
Herald Democrat	Odessa American	Waco Tribune-Herald
Herald and News	Opelika-Auburn News	West Hawaii Today
Herald-Journal	Orange County Register	Winchester Star
Herald-Standard	Orlando Sentinel	Winona Daily News
Houston Chronicle	Peninsula Daily News	Winston-Salem Journal
Houston Defender	Pine Bluff Commercial	Wisconsin State Journal
Hudson Hub-Times	Pittsburgh Post-Gazette	Woodridge Suburban Life
Hunterdon County Democrat	Portland Press Herald	York Daily Record
Imperial Valley Press	Post-Standard	York Dispatch
Independent	Powhatan Today	amNewYork
Jackson Citizen Patriot	Press-Register	
Johnson City Press	Public Opinion	

B Paramter Settings of LSTM Model

Table 4: Parameter Settings of LSTM

Parameters	Value
DROPOUT	0.5
BATCH_SIZE	32
EMBEDDING_SIZE	128
HIDDEN_SIZE	64
NUM_EPOCHS	10

Note: The main parameters ater tuning of the model are the dropout ratio (DROPOUT) which prevents overfitting, the number of data used in each training round (BATCH_SIZE), the number of embedding and hidden layers (EMBEDDING_SIZE and HIDDEN_SIZE), and the number of training rounds (NUM_EPOCHS).

C Summary Statistics

Table 5: Summary Statistics For Variables At Newspaper Level

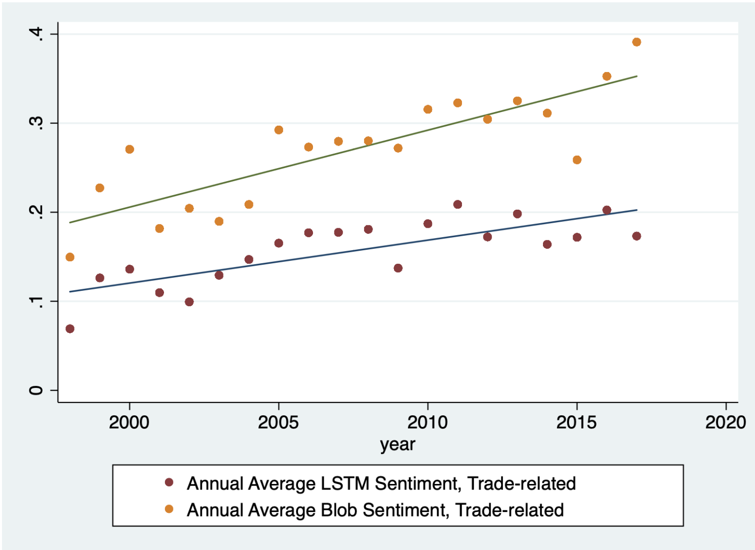
	# Obs.	Mean	Std. Dev.	Min	Max
Panel A: Change of Newspaper Media Slant					
LSTM Measurement					
$\Delta NegRatio$	3,140	0.306	0.289	0.000	1.000
$\Delta NegRatio$: trade-related	3,140	0.157	0.290	0.000	1.000
$\Delta NegRatio$: non-trade-related	3,140	0.488	0.380	0.000	1.000
TextBlob Measurement					
$\Delta NegRatio$	3,140	0.482	0.381	0.000	1.000
$\Delta NegRatio$: trade-related	3,140	0.271	0.399	0.000	1.000
$\Delta NegRatio$: non-trade-related	3,140	0.488	0.380	0.000	1.000
Panel B: Change in Import Exposure and Controls at Newspaper Level					
$\Delta Imports$	3,140	5.256	2.992	0.000	13.201
Population Share of Asian (%)	3,140	0.056	0.044	0.000	0.376
Population Share of Bachelor's Degree (%)	3,140	0.318	0.051	0.000	0.429
Polulation Share of Female (%)	3,140	0.505	0.041	0.000	0.516
Median Income (in ten thousand U.S. dollars)	3,140	3.847	0.546	0.000	4.930

D Appendix

Figure 2: Annual Average Sentiment, 1998-2017

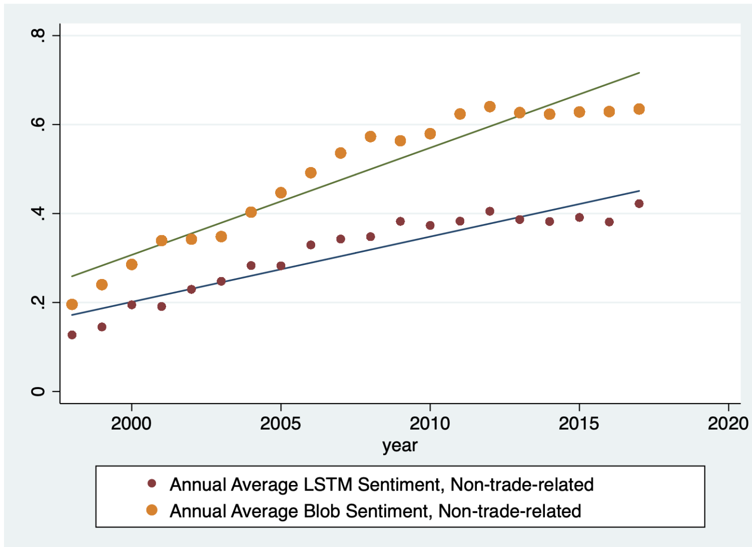
Both annual average negative scores measured by LSTM and Blob in all China-related articles exhibit an upward trend over 1998 to 2017, except for an outlier in the year of 2007. The reportings classified as negative by TextBlob are fewer than that by LSTM.

Figure 3: Annual Average Sentiment on Trade-related Reports, 1998-2017



Both annual average negative scores measured by LSTM and Blob in all trade-related articles on China exhibit an upward trend over 1998 to 2017. The trade-related reportings classified as negative by TextBlob are fewer than that by LSTM.

Figure 4: Annual Average Sentiment on Non-trade-related Reports, 1998-2017



Both annual average negative scores measured by LSTM and Blob in all non-trade-related articles on China exhibit an upward trend over 1998 to 2017. The non-trade-related reportings classified as negative by TextBlob are fewer than that by LSTM. The ratio of negative non-trade-related reportings in all China-relevant articles is generally higher than the negative ratio on trade-related reportings.