



Contents lists available at ScienceDirect

Journal of International Money and Finance

journal homepage: www.elsevier.com/locate/jimfImport competition and U.S. sentiment toward China[☆]Rabah Arezki^{a,*}, Duong Le^b, Ha Nguyen^c, Hieu Nguyen^d^a French National Research Center (CNRS)-Universite Clermont Auvergne, CNRS-IRD-CERDI, Clermont-Ferrand, France^b Fulbright University Vietnam, Ho Chi Minh, Viet Nam^c International Monetary Fund, Washington, DC 20431, USA^d Washington University St. Louis, MO, USA

H I G H L I G H T S

- This paper studies how import competition shapes U.S. local sentiment toward China using a news-based index.
- U.S. sentiment toward China peaked in 2007 before turning negative.
- Communities more exposed to import competition from China have experienced greater deterioration in sentiment.
- Trade-induced U.S. sentiment toward China is broad-based, encompassing political, military, and national security issues.
- The findings suggest that competition over trade may have important geopolitical implications through sentiment of local communities.

A R T I C L E I N F O

JEL classification:

E24
F14
F16
J23
J31
L60
O47
R12
R23

Keywords:

Import competition
Sentiment
Fragmentation

A B S T R A C T

We empirically examine how import competition affects sentiment toward China in local communities in the United States using a news-based index for sentiment. Results are threefold. First, U.S. sentiment toward China peaked in 2007 before turning negative. Second, communities more exposed to import competition from China have experienced greater deterioration in sentiment. Third, the trade-induced U.S. sentiment toward China is broad-based, encompassing political, military, and national security issues. These findings suggest that competition over trade may have important geopolitical implications through sentiment of local communities.

[☆] We thank JaeBin Ahn, Andy Berg, Olivier Blanchard, Mai Dao, Jeffrey Frankel, Mercedes Garcia-Escribano, Gregoire Rota-Graziosi, Tarek Masoud, Flavien Moreau, Rafael Machado Parente, Rick van der Ploeg, and Tarik Youssef and seminar participants at the IMF's Western Hemisphere Department for helpful comments. The findings, interpretations, and conclusions expressed in this paper do not necessarily reflect the views of the World Bank and the International Monetary Fund, the Executive Directors of the World Bank and the IMF or the governments they represent. The World Bank and IMF do not guarantee the accuracy of the data included in this work.

* Corresponding author.

Email addresses: Rabah.AREZKI@uca.fr (R. Arezki), duong.le@fulbright.edu.vn (D. Le), hnguyen7@imf.org (H. Nguyen), h.d.nguyen@wustl.edu (H. Nguyen).

<https://doi.org/10.1016/j.jimonfin.2026.103540>

Available online 13 February 2026

0261-5606/© 2026 Published by Elsevier Ltd.

1. Introduction

The strategic rivalry between the United States and China risks fragmenting the global economy.¹ The rivalry is compounded by the large trade imbalance of the U.S. with China, which has become a major issue.² The distributional implications of the so-called “China shock” on local communities have been well documented (Autor et al., 2013a,b; Pierce and Schott, 2016). The dynamic of sentiment of local communities affected by trade could play an important role in shaping the U.S.-China relationship, arguably the most important of the 21st century. To the best of our knowledge, sentiment of local communities in the U.S. toward China has not been systematically analyzed. This paper fills the gap. Specifically, this paper explores how trade-induced shocks affect sentiment toward China in U.S. local communities.

To do so, we use trade penetration measured by changes in China-originated imports and an index of local news sentiment. In line with the pioneering work by Autor et al. (2013a,b, 2014), we exploit cross-market variation in import exposure from initial differences in industry specialization and instrumenting for U.S. imports using changes in Chinese imports by other high-income countries.³ The increase in imports from China, developed by Autor et al. (2019), is used to examine how the China shock affects sentiment of local U.S. communities for the period from 2000 to 2020. The first half of that period saw the steepest increase in U.S. imports from China following China’s accession to WTO. China’s share of world manufacturing exports surged from 4.8 percent in 2000 to 15.1 percent in 2010. To track sentiment in local communities, we construct a news-based sentiment index using computational text analysis methods. The method extracts sentiment from a database of news articles collected from the top 137 U.S. news publishers (ranked by circulation) in the last four decades, totaling more than 420,000 news articles. By mapping news sentiment to commuting zones and instrumenting trade exposure using Chinese exports to other high-income countries, we provide robust causal evidence of trade-induced sentiment that is persistent and geographically heterogeneous.

Our analysis yields several results. First, U.S. sentiment toward China peaked in 2007 before gradually turning negative. Second, local U.S. communities more exposed to import competition from China have experienced a greater deterioration in sentiment. Third, the trade-induced negative effect on U.S. sentiment toward China is broad-based, encompassing political, military, and national security issues. Together, these findings suggest that competition over trade may have important geopolitical implications through the channel of sentiments of local communities.

This paper is related to a strand of literature that examines the causal link between foreign competition and large and persistent decline in U.S. employment. Local labor markets with more-exposed industries experience greater job losses, greater declines in tax revenues and house prices, larger rises in male idleness and premature mortality (Autor et al., 2013a, 2019; Feler and Senses, 2017). Industries experiencing greater exposure to trade with China have seen higher exit of plants (Bernard et al., 2006), larger contractions in employment (Pierce and Schott, 2016; Acemoglu et al., 2016), and lower income for affected workers (Galle et al., 2023; Caliendo et al., 2019). This paper complements the literature by exploring the consequence of the China shock on U.S. sentiment toward China.

Our paper also contributes to a rapidly expanding literature quantifying attitudes and opinions using news media and examining the predictive power of sentiment on economic activity (see Gentzkow et al. (2019) for a review). An example of the use of dictionary-based sentiment analysis in finance is Tetlock (2007), which considers changes in sentiment extracted from the Wall Street Journal’s “Abreast of the Market” column as a predictor of market fluctuations. Born et al. (2014) find that central bank optimism, derived from Financial Stability Report and governors’ speeches, increases equity prices and reduces volatility. Baker et al. (2016) develop a measure of economic policy uncertainty from the frequency of articles published in 10 leading U.S. newspapers that contain relevant terminologies. Shapiro et al. (2022) show that sentiment shocks have a positive effect on consumption, output, and real interest rates.⁴

Our paper is most closely related to Lu et al. (2018) which examine the effect of the surge in Chinese imports on U.S. media slant against China using 147 U.S. local newspapers for the period running from 1998 to 2012. Lu et al. (2018) find that newspapers in counties facing greater exposure to Chinese imports report more negative news about China. The result is consistent with our findings. However, our paper differs from Lu et al. (2018) in at least three important ways. First, our paper is the first to provide systematic evidence of U.S. sentiment toward China. Specifically, we provide more systematic evidence of U.S. sentiment toward China using a vastly larger and more comprehensive dataset of over 420,000 news articles spanning four decades (from 1979 to 2020), allowing us to trace long-run trends and, more importantly, turning points in public sentiment. Second, our paper constructs a text-based sentiment index utilizing the full text of newspaper articles. Our full textual content analysis of all articles offers a more nuanced and robust measure than the one employed in Lu et al. (2018). That text-based sentiment index contrasts with the much cruder media slant measure in Lu et al. (2018), which is based on arithmetic proportion of China-related newspapers with headlines containing keywords deemed to project a negative image of China. Our more extensive approach allows us to go beyond trade-related narratives and document broader spillover effects on public sentiment across other social and political domains. Third, our paper systematically explores non-trade topics. We find significant spillover effect of trade-induced shocks that encompasses several critical

¹ Together, the U.S. and China account for almost 40 percent of the world’s GDP.

² Data from the U.S. Census Bureau show that the total annual U.S.-China trade volume has increased over 560 percent since China’s accession to the World Trade Organization (WTO) in 2001. The trade volume has increased from \$116 billion in 2000 to \$657 billion in 2021. The value of total imports of the U.S. from China reached \$506 billion in 2021. That is over three times the value of exports from the U.S. to China for the same year.

³ There have been subsequent analyses adopting similar trade-exposure measure to study the effect of trade shocks on public services deliveries and government transfers (Autor et al., 2013a, 2019), as well as tax collections and housing prices (Feler and Senses, 2017).

⁴ Further applications of the text-as-data approach includes FOMC communication (Lucca and Trebbi, 2009; Hansen et al., 2018), flu nowcasting (Ginsberg et al., 2009), corruption in US cities (Saiz and Simonsohn, 2013), or economic reform discussions (Arezki et al., 2020).

domains of public sentiment such as political, military, and national security issues. Overall, these innovations allow us to uncover deeper geopolitical implications of trade competition through the lens of local public sentiment.

The remainder of the paper is organized as follows. Section 2 presents the data. Section 3 describes the empirical strategy. Section 4 presents the main results. Section 5 discusses robustness checks. Section 6 concludes.

2. Data and variable formation

2.1. News-based sentiment

We adopt a text-as-data approach to analyze the content of over 420,000 news articles extracted from the Dow Jones FACTIVA database via the Dow Jones “Data, News and Analytics” Platform.⁵ In addition to the date of publication and the articles’ content, Dow Jones provides several classifiers to navigate article searches. These classifiers include language code, subject code, country/region code (where the article originates from), country discussed by the article, industry code, and publisher information. For our analysis purposes, we restrict the article search and data to cover all articles that included a discussion about China (from the classifier “country discussed by the article”)⁶ and originated from one of the top 137 U.S. news publishers (in terms of circulation).⁷ We obtain 420,642 news articles in total, from 1979 until June 2020. Fig. A-1 shows the number of articles in our dataset over time. The increase in the number of articles discussing China reflects the expansion of FACTIVA and the importance of China in U.S. discussions. Table A-1 shows a tabulation of the number of China-related articles originated from the top 30 U.S. publishers in our database.

To examine the impact of import exposure across dimensions of public attitude, we use FACTIVA’s classification in our main specification. All articles are classified by FACTIVA into one or more of a total of six groups, including (i) *Economics news*, (ii) *Commodity/Financial Markets News*, (iii) *Industry news*, (iv) *Politics/Social News*, (v) *General/Routine News*, and (vi) *Other News*. *Economic news* contains macroeconomic news (for example, “Trump victory freezes trade deals”). *Commodity/Financial Markets news* contains news about markets for trading in financial instruments or commodities (e.g., “Europe, Asia Fall Further; Toronto Rises”). *Industry news* contains all corporate and industrial news (e.g., “CDC Games Hits Record Average Daily Revenue in China”). *Politics/Social news* includes various topics: military and security, human rights and civil liberties, politics and international relations, arts and sciences, socioeconomic news and other social issues (e.g., “Insurer Threatens to Sue Chinese Magazine Over Critical Article”). *General/Routine news* consists of weather reports, sport news or coverage on personal and lifestyles (e.g., “Earthquake Rocks China — Government Faces a Fresh Crisis as Thousands Are Killed”). Articles that don’t belong to any of these groups are marked as *Other news* (e.g., “Business Education: On B-School Test, U.S. Can’t Compete With Asia”). As a robustness check, we employ a machine-learning method to classify news articles into semantic clusters. This approach—Vector Space Semantic methodology (Mikolov et al., 2013)—generates word vector representations by mapping specific vocabulary items in high-dimensional space based on context probabilities (i.e., identifying words that tend to co-occur with a target word or term, and associated frequency). Appendix C describes the machine-learning-based classification approach in more details.

For each news article, we utilize a method popular in the literature to compute a value representing a sentiment index. The index is derived from taking the difference between the number of positive and negative words and scaling it by the sum of positive and negative words in the article.⁸ By design, this index lies in the range $[-1, 1]$, where a higher value indicates more positive public sentiment toward China. A value of 0 indicates a neutral sentiment.⁹

$$\text{sentiment index} = \frac{\text{N}^\circ \text{ positive words} - \text{N}^\circ \text{ negative words}}{\text{N}^\circ \text{ positive words} + \text{N}^\circ \text{ negative words}} \quad (1)$$

Fig. 1 plots the average value of the index of sentiment computed from all U.S. news articles referring to China in the sample period. The U.S. sentiment toward China exhibits distinctive trends during three sub-periods: China’s pre-WTO period (1980–2001), China’s WTO accession (2001–2017), and the contemporary period (2017–2020). Prior to China’s accession to the WTO in 2001, the U.S. sentiment toward China were relatively neutral, with the annual average sentiment index fluctuating around zero for most years, except for a considerable dip in 1989 following the Tiananmen Square incident which led the U.S. government to freeze diplomatic relations with China. China joining the WTO in 2001 was followed by over a decade-long growth surge and rising global importance. This period during which U.S.-China annual trade volume increased over three-fold coincided with positive U.S. sentiment toward China. The sustained period of positive U.S. sentiment toward China ended in 2017, coinciding with the start of the U.S.-China trade war. Continuous sharp drops brought the average annual U.S. sentiment index toward China to the negative territory for the first time since the late 1990s in 2019, followed by the decline to its lowest value during the entire analysis sample recorded in 2020.

China import penetration had been shown to have heterogeneous effects across U.S. localities (Autor et al., 2013a,b, 2014). To capture the spatial variation in sentiment across U.S. localities, we opt for the calculation of public sentiment index using just the content of news articles published by local news sources, and omit articles from well-known national news sources. National sources

⁵ Dow Jones FACTIVA is a global repository of news covering over one billion articles published in many different languages. See <https://www.dowjones.com/professional/factiva/>

⁶ Note that China alone or China and other countries can be discussed in one article. For example, if an article discusses the US-China political relationship, the United States and China are tagged as countries discussed by the article.

⁷ We investigated the extent of cross-ownership of newspapers in our sample. We identified only one source of overlap, which is Hearst Communications, which owns several outlets (e.g., Hearst Newspapers, Boston Herald). The ownership bias is thus limited and unlikely to drive our main results.

⁸ This text-to-data method of deriving sentiment index is widely used in the literature. For examples, see Tetlock (2007) and Shapiro et al. (2022).

⁹ We use the Python package `polyglot` (which builds on the work of Chen and Skiena (2014)) to compute the sentiment value.

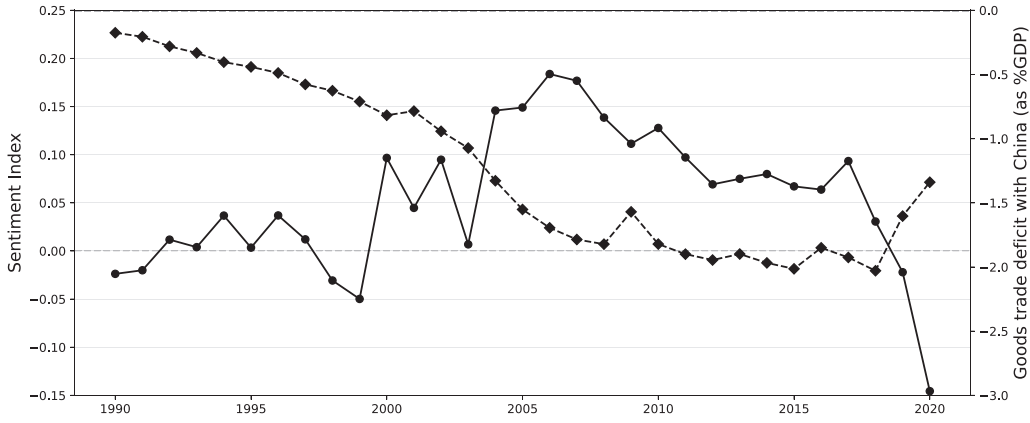


Fig. 1. Average sentiment computed from U.S. news articles referring to China. The solid line (left vertical axis) plots the annual-average sentiment index value computed from all U.S. news articles included in the sample period. News sentiment for 2020 are for the first six months. The dashed line (right vertical axis) plots the U.S.'s good trade deficit with China, as percentage of U.S. GDP.

include ABC News, CBS News, CNN, Fox News, MSNBC, NBC News, NY Times, and Wall Street Journal. These news sources cater to national audiences and are not necessarily representative of sentiment in individual local communities. After removing national news sources, we are left with 338,894 articles from 156 local news sources.¹⁰

2.2. Trade exposure in U.S. localities

2.2.1. Commuting zones

To examine the impact of trade exposure on public sentiment across U.S. localities, we follow the existing literature in defining geographic entities referred to as “commuting zones” (Tolbert et al., 1996; Autor et al., 2013a,b, 2014). Unlike the official administrative divisions (e.g., counties, cities, or towns), a commuting zone groups counties characterized by strong commuting patterns together with a total of 722 zones covering the entire mainland U.S.¹¹ The use of commuting zones as appropriate units of analysis to study the impact of trade shocks on local U.S. labor market has been discussed in earlier studies. By design, a commuting zone better captures the overlap between employment and residential locations of an average individual, even if these locations might belong to different counties. Using commuting zones also helps circumvent issues of selective migration, for instance, migration from cities to suburbs. An analysis at the commuting-zone level, therefore, balances out the need to have a confined geographic unit that collectively captures the impact of trade shocks on labor market, public services, and local sentiment, while still maintaining appropriate spatial granularity.

2.2.2. Measure of exposure to Chinese imports

We use a measure of exposure to Chinese imports in commuting zones as a source of variation in local U.S. sentiment. Following Autor et al. (2019), we use change in import penetration from China to the U.S. in commuting zones as a source of variations:

$$\Delta IP_{it}^{CU} = \sum_j \frac{L_{ij90}}{L_{i90}} \Delta IP_{jt}^{CU} \quad (2)$$

In Eq. (2), the change in import penetration from China to the U.S. in a commuting zone i is the weighted average of the change in import penetration of all industries j , ΔIP_{jt}^{CU} , present in the commuting zone during period τ . The weight is the initial employment share of industry j in commuting zone i , $\frac{L_{ij90}}{L_{i90}}$. L_{ij90} is the initial (i.e., 1990) employment in commuting zone i in sector j , L_{i90} is the initial total employment in commuting zone i . Our paper uses the exposure from Chinese imports from Autor et al. (2019) for the periods between 2000 and 2014 (in the baseline regression), and between 1990 and 2010 (in the extension).

The expression in Eq. (2) apportions the change in the value of U.S. imports from China in a specific industry depending on how employment in that industry is initially distributed across commuting zones in the U.S. and then re-scales this value by total commuting zone employment.

The change in import penetration for each industry j is calculated as the change in the value of U.S. imports from China in industry j during period τ , divided by initial U.S. absorption (which is industry shipment in the U.S. in 1991, Y_{j91} , plus net imports,

¹⁰ The 156 local news sources in our data belong to 119 news publishers. Several publishers own multiple local news publishers. For example, MediaNews Group Inc. is a publisher headquartered in Denver, Colorado. Their publications belong to a total of 8 local news sources, including the Boston Herald headquartered in Massachusetts, The Orange County Register headquartered in California, and the Denver Post headquartered in Colorado. For our analysis, we strictly consider the headquarters of local news sources when identifying news exposure to a U.S. local community. The geographical matching is discussed in details in Section 2.3.

¹¹ Autor et al. (2013a) provides a complete U.S. counties-commuting zones concordance.

$M_{j91} - X_{j91}$):

$$\Delta IP_{jt}^{CU} = \Delta M_{jt}^{CU} / (Y_{j91} + M_{j91} - X_{j91}) \quad (3)$$

For the period from 2000 to 2014, a typical (median) commuting zone has the change in import penetration equal to 0.998. To interpret this value, assuming workers in a commuting zone is entirely employed in an industry X, the change in import penetration of 0.998 implies that the change in import values of industry X between 2000 and 2014 almost equals initial U.S. absorption.

2.3. Mapping news sources to commuting zones

We measure U.S. localities' average sentiment by mapping locations of news sources to each commuting zone. To do so, we first calculate the coordinates of the local headquarters of each news source in our data (156 local news sources) and compute the geographical centre of each commuting zone (722 zones). For each commuting zone, we generate a 300-mile spherical buffer from the centre and match the locations of all news sources that fall within this radius. The 300-mile radius is an arbitrary cutoff to approximate feasible distance for daily distribution of articles' physical prints from the publisher to the commuting zone. In the subsequent analysis, we provide robustness checks to the main result with varying buffer radius.

Once the set of news sources are identified within the commuting zone's buffer, we take the weighted average of sentiment by each new source. The weight is based on the physical distance between the new source's location and the commuting zone's centre to assign the importance of each source toward the computation of commuting zone's public sentiment:

$$\omega(\text{source}_i, \text{czone}_j) = \frac{1 / \text{distance}(\text{source}_i, \text{czone}_j)}{\sum_k 1 / \text{distance}(\text{source}_k, \text{czone}_j)} \quad (4)$$

Overall, 560 out of the total 722 U.S. commuting zones are matched with at least a news source falling within their 300-mile buffer, and thus is assigned a non-missing sentiment index value. Fig. 2 shows the geographic distribution of the local news sources in our database and colors commuting zones by the number of local news sources by which they are influenced. Darker-shaded areas represent local communities influenced by more news sources. It is noted that our news database might not cover the universe of available news sources in the U.S., and in fact, might lack coverage of small local publishers in relatively remote areas (i.e., small publishers not belonging to the list of 137 publishers from which we collected news articles via FACTIVA). Therefore, several commuting zones—mostly low-population localities—do not geographically fall inside the buffer radius of any news sources and are omitted from our analysis. To the extent that small local news sources do not often publish articles discussing China-related content,¹² such omission arguably does not systematically affect our estimates.

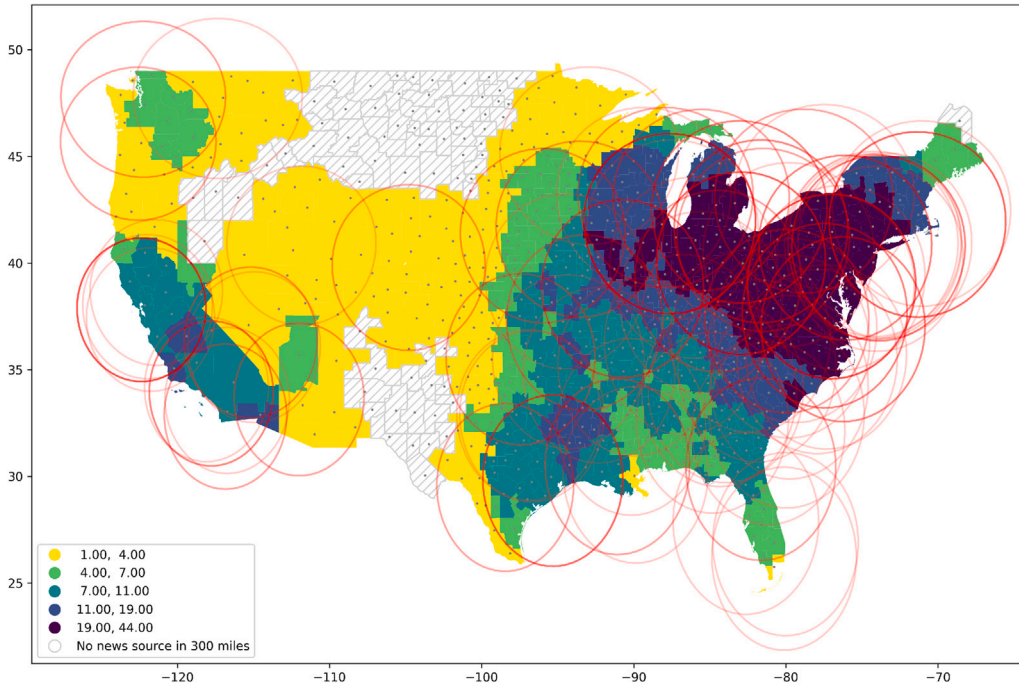


Fig. 2. The distribution of news sources and the number of sources mapped to each commuting zones. Each circle denotes a radius of 300 miles from a news source (156 in total). Each gray dot corresponds to a commuting zone's centre. Color scale represents quantiles of commuting zones based on the number of corresponding sources.

¹² The bottom 10 news publishers in our 137-publisher database only record on average 2.5 articles with a China-related topic during the entire 1979–2020 period.

3. Estimation strategy

3.1. Baseline specification

We estimate the effect of trade exposure on different dimensions of U.S. sentiment of local communities toward China as follows:

$$\Delta Y_{igr} = \beta_0 + \beta_1 \Delta IP_{ir}^{CU} + X_i' \beta_2 + \alpha_g + \varepsilon_{igr} \quad (5)$$

where Y_{igr} is the change in the average news-based sentiment index in commuting zone (CZ) i in topic group g during time interval τ . Our focus is on the period between 2000 and 2014, following China's accession to WTO and fast rising import to the U.S. originating from China. The main explanatory variable is the corresponding change in CZ-level import penetration from China to the U.S. ΔIP_{ir}^{CU} , instrumented by China's exports to other advanced countries, ΔIP_{ir}^{CO} . X_i denotes locality controls capturing time-invariant demographic characteristics and labor force composition in the year 2000. Additionally, we control for semantic group dummies α_g for news articles. The standard errors ε_{igr} are clustered at the state level.

One concern in identifying a causal relationship between U.S. imports of Chinese goods and local sentiment across U.S. commuting zones is the potential existence of confounding factors. These factors, such as the (spatially) heterogeneous changes in U.S. income, could drive both sentiment and changes in U.S. imports. To isolate the exogenous variation in U.S. imports from China, following Autor et al. (2019), we use the change in other high-income country imports from China, ΔIP_{ir}^{CO} , as an instrument for ΔIP_{ir}^{CU} :

$$\Delta IP_{ir}^{CO} = \sum_j \frac{L_{ij80}}{L_{i80}} \Delta IP_{jr}^{CO} \quad (6)$$

4. Main results

4.1. Baseline results

Table 1 reports the baseline results using our estimation strategy including the instrument variable described in Eq. (6). Overall, we find a significant negative causal effect of China-import penetration on U.S. sentiment during the period from 2000 to 2014. The estimated coefficients suggest over 0.03 unit decline in sentiment toward China in a commuting zone experiencing one-unit larger import penetration, which is about the median change between 2000 and 2014. Our baseline result complements existing findings in the literature about the distributional implications of trade exposure to China. Recall that the aggregate sentiment moved between -0.15 and 0.20 (Fig. 1), therefore a 0.03 unit decline in sentiment is quite meaningful.

4.2. Heterogeneity by topics

Having shown an aggregate negative effect of Chinese imports pressure on sentiment, we now explore the granularity of sentiment. A natural question is whether import penetration has a heterogeneous effect across different themes of the U.S.-China relationship. We rely on FACTIVA's classification of news articles by themes and generate 6 groups including *Economic news*, *Commodity/financial-markets news*, *Industry news*, *Political/social news*, *General/routine news*, and *Other news*. The heterogeneity in trade-induced effect on

Table 1
Impact of China-import competition on U.S. public sentiment (Baseline).

IV regression	(1)	(2)	(3)
	Δ Average sentiment (2000–2014)		
Δ Import penetration (2000–2014)	−0.0180* (0.00948)	−0.0191** (0.00889)	−0.0213** (0.00992)
Constant	✓	✓	✓
Observations	2521	2521	2521
R-squared	0.054	0.061	0.062
Topic group FE	✓	✓	✓
Occupational controls	✓	✓	✓
Race controls		✓	✓
Demographic controls			✓
Number of commuting zones	560	560	560
Standardized coefficient	−1.8987	−2.1485	−2.1472

Note: Observation for 6 news-topic groups across 560 commuting zones. Dependent variable: change in average sentiment index for each group between 2000 and 2014. News sources are mapped to commuting zones based on physical proximity and volume of published articles. Occupational controls include start-of-period indices of employment in routine occupations, employment share of manufacturing, and employment in offshorable occupations. Race controls are dummies representing start-of-period shares of commuting zone population that is Hispanic, Black, Asian, and other races. Demographic controls include start-of-period shares of foreign born population, college educated, and female employment. The list of control variables follows Autor et al. (2019). Reported standardized coefficients are for the main explanatory variable. Robust standard errors in parentheses are clustered at the state level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

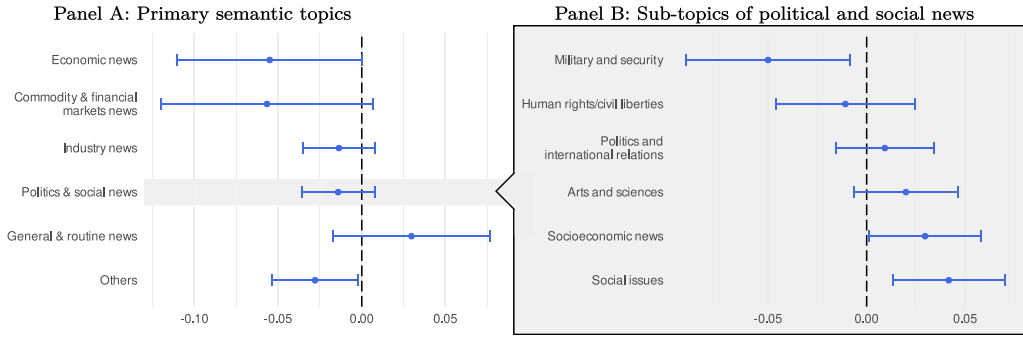


Fig. 3. Heterogeneous effects of China-import penetration across themes. Panel A shows the 90% confidence intervals of the effect of import competition from China on changes in local sentiment. Panel B shows similar statistics but only for the sub-topics in political/social news. The dependent variable is the change in sentiment between 2000 and 2014. News sources are mapped to commuting zones based on physical proximity and volume of published articles. Occupational controls include start-of-period indices of employment in routine occupations, employment share of manufacturing, and of employment in offshorable occupations. Race controls are dummies representing start-of-period shares of commuting zone population that is Hispanic, Black, Asian, and other races. Demographic controls include start-of-period shares of foreign-born population, college educated, and female employment. The list of control variables follows Autor et al. (2019). Robust standard errors in parentheses clustered by state.

sentiment can be estimated with an interaction term between each of the themes and the Chinese imports indicator:

$$\Delta Y_{igr} = \beta_0 + \beta_1 \Delta IP_{ir}^{CU} + X_i' \beta_2 + \beta_3 \Delta IP_{ir}^{CU} \times \alpha_g + \alpha_g + \varepsilon_{igr} \quad (7)$$

where the new term α_g is an indicator for each theme. β_3 is the estimated coefficient corresponding to the interaction term $IP_{ir}^{CU} \times \alpha_g$ and illustrates the impact of Chinese imports on local sentiment about China across different themes.

Fig. 3 plots the point estimates and 90-percent confidence intervals of the effect of Chinese imports on U.S. sentiment toward China (i.e., $\beta_1 + \beta_3$) for each of the six semantic themes. By design, a negative estimate associated with a theme suggests that higher import competition from China reduces local U.S. sentiment toward China based on news articles categorized into that thematic news group. The estimates in Panel A suggest that exposure to Chinese imports drives local U.S. sentiment about China downward across all thematic groups except for *General and routine news*. The largest and most statistically significant effects are observed for articles grouped into *Economic news* and *Commodity & financial-market news*.

It is not surprising to observe a strong and negative effect on U.S. sentiment expressed through news articles on topics directly related to themes such as economics, financial markets, and industry. However, the fact that the marginal trade-induced effect is also negative for themes such as political and social news topics suggests that pressure from import competition could also have important implication across “non-economic” dimensions of sentiment.

To further investigate the potential spillover effect on sentiment expressed in political and social news, we examine a more granular set of topics within these themes and estimate the effect of China trade-induced shock on sentiment. Among the available sub-topics, we find negative estimates for articles in *Military & national security*, and—albeit insignificant at conventional confidence intervals—*Human rights & civil liberties*. The negative trade-induced effect on sentiment expressed in articles discussing national security and human rights suggest trade-induced shocks could influence the U.S. awareness and attitude toward the most politically salient issues concerning China in contemporary time.

To explore the linkage between the penetrated industries and news themes, we have borrowed checks from Goldsmith-Pinkham et al. (2020), specifically adopting the Rotemberg weight over each news theme. Results of the decomposition presented in Fig. D-1 in Appendix D show the heterogeneous effects of different industries on news sentiment. Consistent with the results obtained by Goldsmith-Pinkham et al. (2020), we find that the weight of penetrated industries to be heavily skewed. Indeed, the top five industry-level instruments across news themes are consistently: *Radio + TV Communications Equipment*; *Furniture and Fixtures*; *Women’s Non-Athletic Footwear*; *Women’s and Children’s Underwear*; and *Men’s + Boys’ Underwear + Nightwear*. Indeed, these five industries receive about half of the absolute weight in the various coefficients associated with news sentiment.

Notwithstanding the recurrence of these five industries across types of news, there are subtle differences in terms of signs and relative contributions of these penetrated industries. In the below, we provide some highlights on how the analysis of Rotemberg weights allows us to gain new insights into the heterogeneous effect of penetrated industries on sentiment relative to various news themes. Notably, the weight associated with the “Furniture and Fixtures” industry is by far the most important weight and is negatively associated with the coefficient of “Commodity & Financial Markets” news theme. The sign associated with this specific penetrated industry weight reflect the depth of the import competition driving down US sentiment toward China on news theme related to “Commodity & Financial Markets”. Add to that the central role of China in international commodity markets coupled with the fact that the “Furniture and Fixtures” industry is commodity based.

When it comes to the coefficient associated with “Economic news” theme, the “Radio and TV Communication Equipment” industry contributes the most in terms of weight with a positive sign. The sign associated with the weight for the “Furniture and Fixtures” industry is also positive. That is not surprising considering the “Economic news” theme mostly covers macroeconomic news in FACTIVA, a news research tool. In other words, the decomposition suggests that positive developments in these industries where

Table 2
Persistent effect of Chinese imports penetration on U.S. public sentiment.

IV regression	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Δ Average sentiment between 2000 and						
	2014	2015	2016	2017	2018	2019	2020
Δ Import penetration (2000–2014)	−0.0213** (0.00992)	−0.0175* (0.0103)	−0.0188** (0.00942)	−0.0217** (0.00888)	−0.00801 (0.00940)	−0.0153* (0.00836)	−0.0232** (0.00996)
Constant	✓	✓	✓	✓	✓	✓	✓
Observations	2521	2682	2617	2475	2317	2504	2598
R-squared	0.062	0.290	0.225	0.133	0.157	0.231	0.161
Topic group FE	✓	✓	✓	✓	✓	✓	✓
Occupational controls	✓	✓	✓	✓	✓	✓	✓
Race controls	✓	✓	✓	✓	✓	✓	✓
Demographic controls	✓	✓	✓	✓	✓	✓	✓
N° commuting zones	560	560	560	526	560	528	560
Standardized coefficient	−2.1472	−1.6990	−1.9958	−2.4437	−0.8521	−1.8301	−2.3293

Note: Observation for 6 topics groups across 560 commuting zones. Dependent variable: average sentiment indices for each semantic group in 2014, 2015, 2016, 2017, 2018, 2019 and 2020, all subtracted by average sentiment indices for each semantic group in 2000. News sources are mapped to commuting zones based on physical proximity and volume of published articles. Occupational controls include start-of-period indices of employment in routine occupations, employment share of manufacturing, and employment in offshorable occupations. Race controls are dummies representing start-of-period shares of commuting zone population that is Hispanic, Black, Asian, and other races. Demographic controls include start-of-period shares of foreign-born population, college educated, and female employment. The list of control variables follows Autor et al. (2019). Reported standardized coefficients are for the main explanatory variable. Robust standard errors in parentheses are clustered at the state level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

China's penetration happen to be high are associated with positive US sentiment toward China alongside positive macroeconomic developments.

The decomposition of the role of penetrated industries in “Political news” also feature dominant weights from both “TV and Radio equipment” and “Furniture and Fixtures” industries. The former industry is associated with a positive sign while the latter is associated with a negative sign—albeit the difference in absolute values are relatively small. The difference in sign likely reflects that the US sentiment toward China related to “Political news” is differentiated depending on whether the penetrated industry support the service vs. the manufacturing sectors. The decomposition of the role of penetrated industries in “Routine news” is mostly associated with negative signs for all top industries. That reflects the nature of these news themes which covers developments in the penetrated industries with potential negative socioeconomic elements. The signs associated with the weights of the different penetrated industries for “Other news” theme are also mostly negative albeit this news category is less readily interpretable considering its residual nature.

4.3. Persistent effects of import exposure on news sentiment

How persistent is the effect on sentiment? We examine this question by utilizing the news-based sentiment index calculated from news articles published in the years after U.S.-China trade competition bottomed out in the early 2010s. To analyze the effect of sentiment on an annual basis, we derive the outcome variable as the difference between the average sentiment for each year from 2014 to 2019 and the baseline sentiment level in 2000. The main explanatory variable is still the instrumented change in U.S. imports of Chinese goods between 2000 and 2014, from Autor et al. (2019).

The estimation reported in Table 2 suggests that trade competition continued to impact sentiment in the communities exposed to Chinese imports in all the years after the 2000–2014 period. The negative impact on attitude reflected in U.S. news articles that featured China related issues was particularly strong during 2015–2017, and again in 2019 and 2020.

As an additional robustness check, Table B-6 in Appendix B presents the results of Table 2 for an estimation using a balance sample across all specifications from (1) to (7). Missing theme-specific sentiment observations in certain commuting zone-years were linearly imputed. The results indicate that our main estimates are robust to using a balance sample.

Further, we split the sample into two sub-sample periods, namely 2000–2007 and 2008–2014. Results are presented in Table B-1 in Appendix B. Overall, our main results are robust to the sample split, that is, import competition induced a deterioration of news-based sentiment. The results provide new insights into how the trade shock affects sentiment before and after the peak. Quantitatively, standardized coefficients indicate that the effect of import competition from China is bigger for the first sub-period running from 2000 to 2007—corresponding to the rapid increase in China net exports to the US—relative to the second sub-period running from 2007 to 2014—corresponding to a plateau in terms of China net exports.

5. Extension and robustness checks

5.1. Results for the period from 1990–2000

In this extension, we estimate the impact of China-import penetration on U.S. public sentiment using changes in import competition in the decade between 1990 and 2000. This decade experienced the commencing of the China shock in 1992, when China expanded

its “reform and opening” agenda to foster foreign trade and investment, and before the accession of China to the WTO in 2001.

$$\Delta Y_{igt0} = \beta_0 + \beta_1 \Delta IP_{igt0}^{CU} + X'_i \beta_2 + \alpha_g + \varepsilon_{igt0} \quad (8)$$

Results on the effect of trade exposure on U.S. sentiment during 1990–2000 remain significant and negative. The regression coefficients, shown in [Table B-2](#), suggest that the size of the effect could be relatively larger than that of our baseline results shown in [Tables 1](#) and [2](#)—covering the 2000–2014 period. This finding suggests commuting zones more exposed to rising imports from China saw a larger decline or a smaller rise in sentiment toward China during 1990–2000. Note that compared with our baseline regressions for the 2000–2014 period, our dataset contains fewer news articles between 1990–2000 and thus covers fewer commuting zones. Between 1990–2000, our data contain 57,164 articles and cover 431 commuting zones. In comparison, our data for the 2000–2014 period contain 221,980 articles, covering 560 commuting zones.

Further, we have produced placebo tests in [Table B-3](#) in [Appendix B](#). Columns (1) to (3) in the table show the effect of import penetration for the period spanning from 2000 to 2014 on sentiment for the period spanning from 1990 to 2000. In Columns (4) to (6), we show the effect of import penetration from 1990 to 2000 on sentiment for the period spanning from 2000 to 2014. As expected, results point to the absence of placebo effect of future import penetration on past sentiment as shown in Columns (1) to (3). However, there is evidence that past import penetration on future sentiment signals persistence in the effect of import penetration on sentiment. Overall, this useful placebo test confirms that future import penetration does not explain past sentiment.

5.2. Varying physical proximity between news sources and commuting zones

Next, we check the sensitivity of the causal relationship to our method of measuring the degree of news influence in each commuting zone. To do so, we re-estimate the effect using two alternative commuting zone’s proximity buffer radius taking values of 100 and 200 miles (compared to the original 300-mile buffer). Under more restricted sample sizes, the estimated results in [Table B-4](#) show that the trade-induced effect on sentiment is stronger and statistically significant across all specifications, although the number of commuting zones with sentiment data are much smaller.¹³

5.3. Machine-learning topic selection

Lastly, we check for the sensitivity of our estimates to methods in selecting thematic groupings of news articles. Our baseline specification uses the manual topic tags provided by FACTIVA to group news articles into six distinct themes. In this robustness check, we utilize a machine-learning method to systematically classify articles into topical groups instead (see [Appendix C](#) for details). One input parameter of the method is the number of semantic clusters. We allow the value of this input parameter to take three different values of six, eight, and ten topic groups. [Table B-5](#) shows that estimations using these different number of thematic groups yields similar results to those obtained using our baseline specification.

6. Conclusion

This paper examines how import competition affect sentiment toward China in local communities in the United States using a news-based index for sentiment. We find that commuting zones which bear the brunt of the trade shock exhibit significantly more negative shifts in sentiment. This adverse effect is not confined to economic and financial news, but permeates a range of social topics, including personal finance, military/security, and human rights. The negative effect on news-based sentiment persists over time and is robust to various checks. Our findings suggest that competition over trade may have important geopolitical implications through sentiment of local communities.

CRedit authorship contribution statement

Rabah Arezki: Writing – review & editing, Writing – original draft, Formal analysis, Conceptualization. **Duong Le:** Writing – review & editing, Writing – original draft, Formal analysis, Conceptualization. **Ha Nguyen:** Writing – review & editing, Writing – original draft, Formal analysis, Conceptualization. **Hieu Nguyen:** Writing – review & editing, Writing – original draft, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

¹³ There are 384 and 150 CZs matched with at least one nearby local news sources within 200-mile and 100-mile buffer radius, respectively. In comparison, there are 560 CZs matched under the baseline 300-mile radius buffer.

Appendix A. Composition of news in factiva

A.1. News articles mentioning China

Table A-1

News articles about China by top U.S. publishers in FACTIVE database, 1979–2020.

Publisher name	N° articles referring China
Dow Jones & Company, Inc.	85 178
Business Wire, Inc.	57 684
The New York Times Company	29 657
Hearst Communications, Inc.	23 211
Washington Post	21 047
New York Times Digital (Full Text)	16 582
Dow Jones & Company	13 395
Seattle Times	12 509
Investor's Business Daily	11 892
News World Communications, Inc.	10 409
Gannett Co., Inc.	9 129
St. Louis Post-Dispatch	6 123
CQ-Roll Call, Inc.	6 067
National Public Radio, Inc.	5 450
NewsBank Inc.	4 914
MediaNews Group Inc.	4 628
Knight Ridder Digital	4 374
Journal Communications Inc.	3 034
Media General, Inc.	3 007
Grand Rapids Press	2 597
Hollinger	2 410
Deseret News Publishing Co.	2 323
N.Y.P. Holdings, Inc.	2 308
Advance Publications, Inc.	2 184
Boston Herald Library	2 105
Hearst Newspapers	1 980
Union-Tribune Publishing Company	1 873

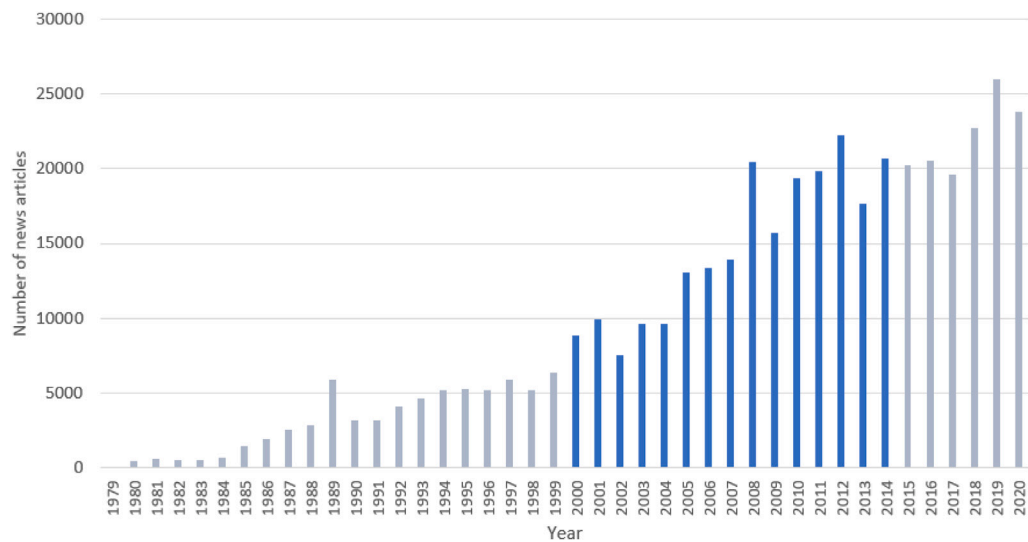


Fig. A-1. Number of news articles in the dataset. Data for 2020 are for the first six months. Our baseline analyses focus on 2000 to 2014.

A.2. News articles mentioning other emerging countries

see Fig. A-2.

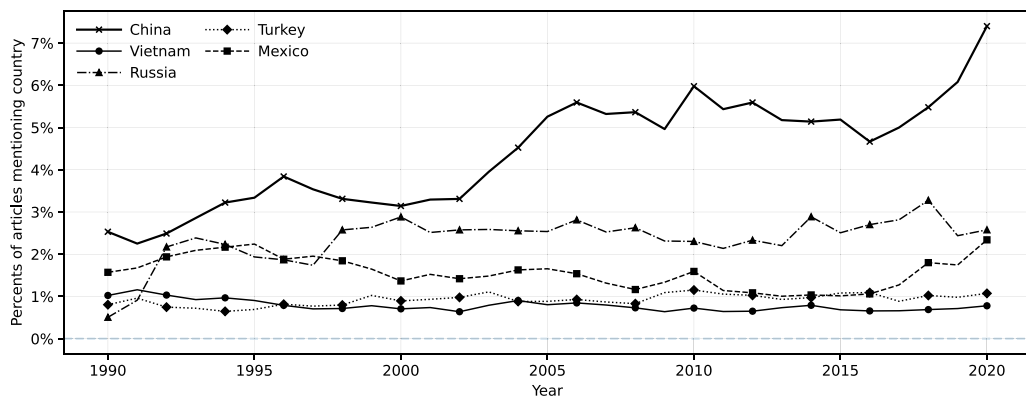


Fig. A-2. Each line plots the share of U.S. news articles mentioning one of the emerging countries.

Appendix B. Extensions and robustness checks

B.1. Regression results before and after 2007

Table B-1
Impact of China-import competition on U.S. public sentiment.

	(1)	(2)	(3)	(4)	(5)	(6)
IV regression	Δ Average sentiment between 2000–2007			Δ Average sentiment between 2007–2014		
Δ Import penetration (2000–2007)	−0.00923* (0.00521)	−0.00962** (0.00470)	−0.0112** (0.00553)			
Δ Import penetration (2007–2014)				−0.0210 (0.0134)	−0.0228* (0.0133)	−0.0233* (0.0136)
Constant	✓	✓	✓	✓	✓	✓
Observations	2521	2521	2521	2521	2521	2521
R-squared	0.053	0.060	0.062	0.050	0.057	0.058
Occupational controls	✓	✓	✓	✓	✓	✓
Race controls		✓	✓		✓	✓
Demographic controls			✓			✓
N° commuting zones	560	560	560	560	560	560
Standardized coefficient	−1.7716	−2.0468	−2.0253	−1.5672	−1.7143	−1.7132

Note: Observations for 6 topics groups across 560 commuting zones. Dependent variable: change in average sentiment index for each group between 2000–2007 (columns 1–3) and between 2007–2014 (columns 4–6). News sources are mapped to commuting zones based on physical proximity and volume of published articles. Occupational controls include start-of-period indices of employment in routine occupations, employment share of manufacturing, and employment in offshorable occupations. Race controls are dummies representing start-of-period shares of commuting zone population that is Hispanic, Black, Asian, and other races. The list of control variables follows Autor et al. (2019). Reported standardized coefficients are for the main explanatory variable. Robust standard errors in parentheses are clustered at the state level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

B.2. Regression results for the 1990–2000 period

Table B-2
Impact of Import Competition from China on U.S. Sentiment (1990–2000).

IV regression	(1)	(2)	(3)
	Δ Average sentiment (1990–2000)		
Δ Import penetration (1990–2000)	–0.0318* (0.0183)	–0.0332* (0.0174)	–0.0312** (0.0156)
Constant	✓	✓	✓
Observations	1271	1271	1271
R-squared	0.174	0.178	0.183
Topic group FE	✓	✓	✓
Occupational controls	✓	✓	✓
Race controls		✓	✓
Demographic controls			✓
Number of commuting zones	431	431	431
Standardized coefficient	–1.7377	–1.9080	–2.0000

Note: Observation for 6 news-topic groups across 431 commuting zones. Dependent variable: change in average sentiment index for each group between 1990 and 2000. News sources are mapped to commuting zones based on physical proximity and volume of published articles. Occupational controls include start-of-period indices of employment in routine occupations, employment share of manufacturing, and employment in offshorable occupations. Race controls are dummies representing start-of-period shares of commuting zone population that is Hispanic, Black, Asian, and other races. Demographic controls include start-of-period shares of foreign born population, college educated, and female employment. The list of control variables follows Autor et al. (2019). Reported standardized coefficients are for the main explanatory variable. Robust standard errors in parentheses are clustered at the state level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

B.3. Placebo tests

Table B-3
Placebo test.

IV regression	(1)	(2)	(3)	(4)	(5)	(6)
	Δ Average sentiment between 1990–2000			Δ Average sentiment between 2000–2014		
Δ Import penetration (2000–2014)	0.00368 (0.0145)	0.00293 (0.0143)	0.00286 (0.0107)			
Δ Import penetration (1990–2000)				–0.0325* (0.0190)	–0.0285* (0.0154)	–0.0359* (0.0185)
Constant	✓	✓	✓	✓	✓	✓
Observations	1271	1271	1271	2521	2521	2521
R-squared	0.179	0.184	0.187	0.044	0.054	0.053
Occupational controls	✓	✓	✓	✓	✓	✓
Race controls		✓	✓		✓	✓
Demographic controls			✓			✓
N° commuting zones	431	431	431	560	560	560
Standardized coefficient	0.2538	0.2049	0.2673	–1.7105	–1.8507	–1.9405

Note: Observation for 6 topics groups across 560 commuting zones. Dependent variable: change in average sentiment index for each group between 1990–2000 (columns 1-3) and between 2000–2014 (columns 4-6). News sources are mapped to commuting zones based on physical proximity and volume of published articles. Occupational controls include start-of-period indices of employment in routine occupations, employment share of manufacturing, and employment in offshorable occupations. Race controls are dummies representing start-of-period shares of commuting zone population that is Hispanic, Black, Asian, and other races. The list of control variables follows Autor et al. (2019). Reported standardized coefficients are for the main explanatory variable. Robust standard errors in parentheses are clustered at the state level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

B.4. Varying news buffer radius

Table B-4

Robustness check with varying buffer zones.

IV regression	(1)	(2)	(3)	(4)	(5)	(6)
Δ Average sentiment (2000–2014)						
Δ Import penetration (2000–2014)	−0.0564* (0.0306)	−0.0596** (0.0281)	−0.0586** (0.0279)	−0.0375** (0.0174)	−0.0339** (0.0164)	−0.0413** (0.0181)
Constant	✓	✓	✓	✓	✓	✓
Observations	634	634	634	1608	1608	1608
R-squared	0.088	0.106	0.119	0.059	0.076	0.080
Proximity buffer radius	100-mile	100-mile	100-mile	200-mile	200-mile	200-mile
Topic group FE	✓	✓	✓	✓	✓	✓
Occupational controls	✓	✓	✓	✓	✓	✓
Race controls		✓	✓		✓	✓
Demographic controls			✓			✓
Number of commuting zones	150	150	150	384	384	384
Standardized coefficient	−1.8431	−2.1210	−2.1004	−2.1552	−2.0671	−2.2818

Note: Observation for 6 topics groups across 150 commutizones for models (1–3) and 384 commuting zones for models (4–6). Dependent variable: change in average sentiment index for each group between 2000 and 2014. News sources are mapped to commuting zones based on physical proximity and volume of published articles. Occupational controls include start-of-period indices of employment in routine occupations, employment share of manufacturing, and employment in offshorable occupations. Race controls are dummies representing start-of-period shares of commuting zone population that is Hispanic, Black, Asian, and other races. Demographic controls include start-of-period shares of foreign born population, college educated, and female employment. The list of control variables follows Autor et al. (2019). Reported standardized coefficients are for the main explanatory variable. Robust standard errors in parentheses are clustered at the state level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

B.5. Using machine learning to identify topic clusters

Table B-5

Impact of China-import competition on U.S. public sentiment.

IV regression	(6 groups)	(8 groups)	(10 groups)
Δ Average sentiment (2000–2014)			
Δ Import penetration (2000–2014)	−0.0201** (0.00837)	−0.0187* (0.0110)	−0.0123 (0.00885)
Constant	✓	✓	✓
Observations	2768	3507	4445
R-squared	0.159	0.114	0.103
Controls	✓	✓	✓
Topic FE	✓	✓	✓
Number of CZs	560	560	560
Standardized coefficient	−2.4014	−1.7000	−1.3898

Note: Observation for 6, 8 and 10 semantic groups across 560 commuting zones (columns 1, 2, and 3 respectively). Dependent variable: change in average sentiment index for each group between 2000 and 2014. News sources are mapped to commuting zones based on physical proximity and volume of published articles. Occupational controls include start-of-period indices of employment in routine occupations, employment share of manufacturing, and employment in offshorable occupations. Race controls are dummies representing start-of-period shares of commuting zone population that is Hispanic, Black, Asian, and other races. Demographic controls include start-of-period shares of foreign born population, college educated, and female employment. The list of control variables follows Autor et al. (2019). Reported standardized coefficients are for the main explanatory variable. Robust standard errors in parentheses are clustered at the state level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

B.6. Persistence analysis with balanced sample

Table B-6

Persistent effect of China imports penetration on U.S. public sentiment.

IV regression	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Δ Average sentiment between 2000 and						
	2014	2015	2016	2017	2018	2019	2020
Δ Import penetration (2000–2014)	−0.0213** (0.00992)	−0.00966 (0.00909)	−0.0175* (0.00913)	−0.0184** (0.00881)	−0.00957 (0.00781)	−0.0105 (0.00716)	−0.0224** (0.00879)
Constant	✓	✓	✓	✓	✓	✓	✓
Observations	2521	2521	2521	2521	2521	2521	2521
R-squared	0.062	0.272	0.202	0.085	0.174	0.172	0.102
Topic group FE	✓	✓	✓	✓	✓	✓	✓
Occupational controls	✓	✓	✓	✓	✓	✓	✓
Race controls	✓	✓	✓	✓	✓	✓	✓
Demographic controls	✓	✓	✓	✓	✓	✓	✓
N° commuting zones	560	560	560	526	560	528	560
Standardized coefficient	−2.1472	−1.0627	−1.9168	−2.0885	−1.2254	−1.4665	−2.5483

Note: Observation for 6 topics groups across 560 commuting zones. Dependent variable: average sentiment indices for each semantic group in 2014, 2015, 2016, 2017, 2018, 2019 and 2020, all subtracted by average sentiment indices for each semantic group in 2000. Sentiment changes from 2015 onwards are linearly imputed so that they match the 2014 sample. News sources are mapped to commuting zones based on physical proximity and volume of published articles. Occupational controls include start-of-period indices of employment in routine occupations, employment share of manufacturing, and employment in offshorable occupations. Race controls are dummies representing start-of-period shares of commuting zone population that is Hispanic, Black, Asian, and other races. Demographic controls include start-of-period shares of foreign-born population, college educated, and female employment. The list of control variables follows Autor et al. (2019). Reported standardized coefficients are for the main explanatory variable. Robust standard errors in parentheses are clustered at the state level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Appendix C. Topic analysis of news articles

We created word vector representations (Mikolov et al., 2013) to represent words through vectors relying on Vector Space Semantics methodology.¹⁴ This is done by mapping specific vocabulary items in high-dimensional space based on contextual probabilities (i.e., identifying words that tend to co-occur with a target word or term, and how often).

Vector space representations have been shown to efficiently summarize the thematic relationships between words in a corpus and enable to measure thematic relatedness between any two given words. Word-vector models can computationally determine “semantic clusters” containing words that belong together.

We used a pre-trained model from Google that includes word vectors for a vocabulary of 3 million words and phrases which they trained on roughly 100 billion words from a Google News dataset. Typical vector dimensionality used in implementations is between 100 and 300. In our implementation, the vector dimensionality is 300. Vector representations of words were computed using the package *gensim* (Rehurek and Sojka, 2011) in Python.

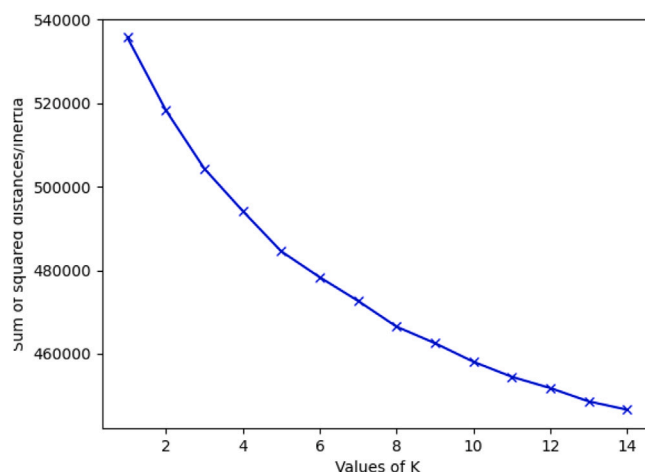
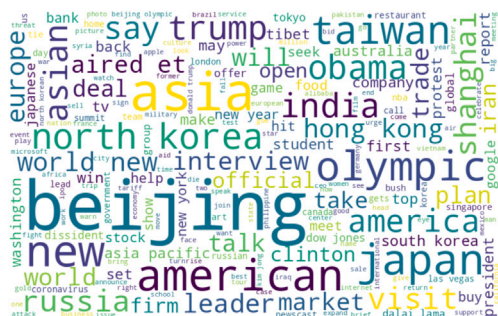


Fig. C-1. Sum of squared distances to each cluster centers by the number of cluster (K).

¹⁴ Specifically, we use the word2vec family of vector space models.

Using this model, each article’s title in our dataset is mapped into the 300 dimensions of vector representation. We then proceed to use the K-means algorithm to cluster these titles into semantic clusters. We implemented it with different number of clusters (K). Higher number of clusters means that clusters are better fitted (lower sum of squared distance to clusters’ centroids) and display more nuanced distinction in topics that articles represent. However, the more clusters we used, the more fragmented our data becomes, and inference of clusters’ themes are increasingly difficult (Figs. C-1 and C-2).

(a) General news



(b) News aggregation



(c) Reports and researches



(d) News briefings



(e) Market & technology



(f) Social, culture, and sports



(g) Human rights



(h) Politics



Fig. C-2. Word clouds for each topic clusters.

Appendix D. Rotemberg decomposition

To study the heterogeneous effects of different industries on news sentiment, we deployed the Rotemberg decomposition method as demonstrated by Goldsmith-Pinkham et al. (2020). The results of these decompositions are shown in Fig. D-1.

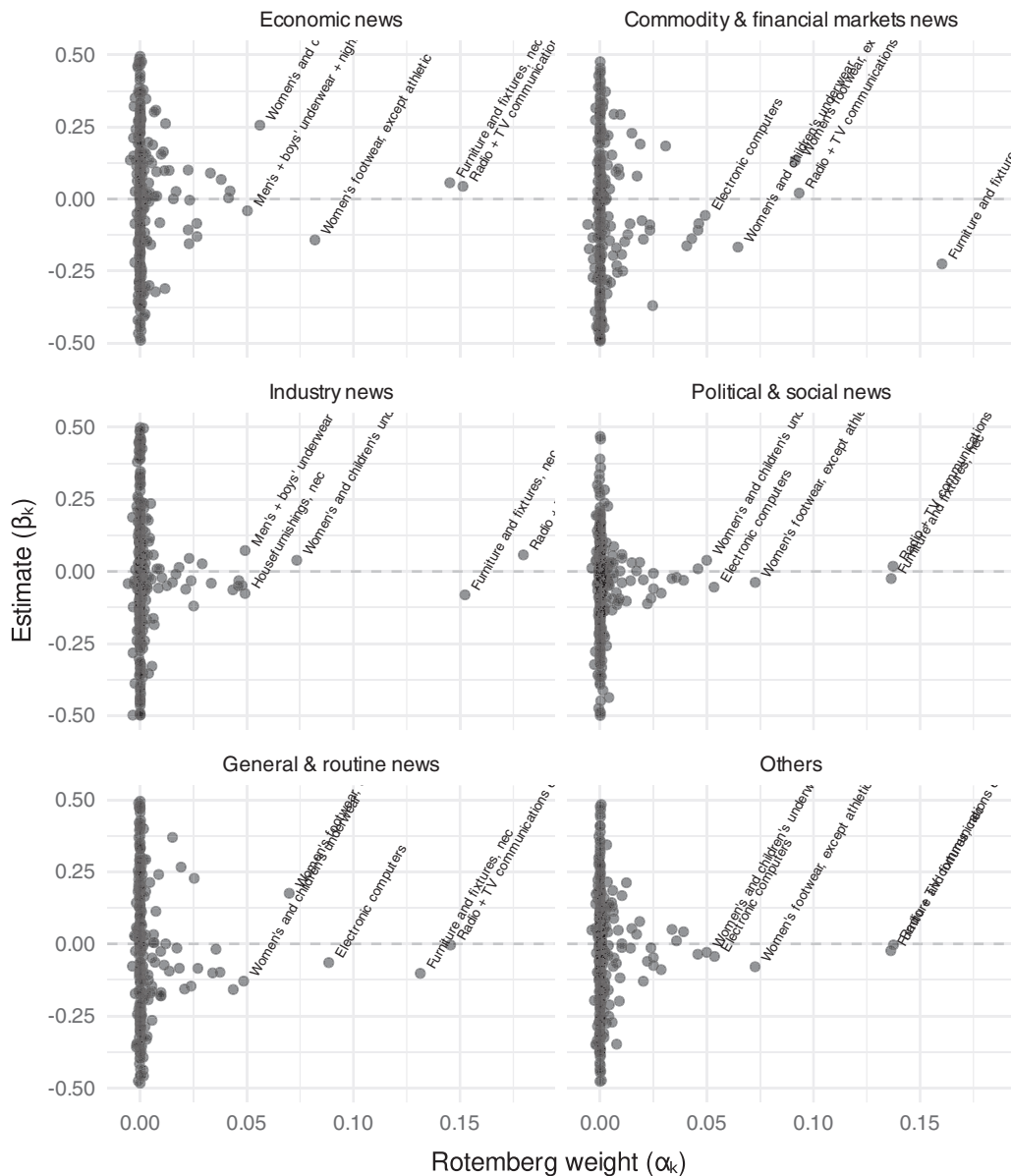


Fig. D-1. Rotemberg decomposition of main instrument. This graph plots on the horizontal axis the Rotemberg contribution for each industry-level instrument and on the vertical axis their corresponding estimates. Each of the six panels represents (separately) the decomposition for each news group. Labels are shown for the top 5 industries (as represented by their Rotemberg weight) in each news group. The sum of all Rotemberg weight adds up to one.

Similar to Goldsmith-Pinkham et al. (2020), we find industry impacts to be heavily skewed. The top 5 industry-level instruments (which consistently features *Radio + TV communications equipment*, *Electronic computers*, *Furniture and fixtures*, *Women's non-athletic footwear*, *Women's and children's underwear*, and *Men's + boys' underwear + nightwear*) receive about half of the absolute weight in the estimator.

Data availability

Data will be made available on request.

References

- Acemoglu, D., Autor, D., Dorn, D., Hanson, G.H., Price, B., Jan 2016. Import competition and the great US employment sag of the 2000s. *J. Labor Econ.* 34 (S1), S141–S198.
- Arezki, R., Djankov, S., Nguyen, H., Yotzov, I., 2020. Reform Chatter and Democracy, World Bank Policy Research Working Paper 9319.
- Autor, D., Dorn, D., Hanson, G., Sep 2019. When work disappears: manufacturing decline and the falling marriage market value of Young Men. *Am. Econ. Rev.: Insights* 1 (2), 161–178.
- Autor, D.H., Dorn, D., Hanson, G.H., Oct 2013. The China syndrome: local labor market effects of import competition in the United States. *Am. Econ. Rev.* 103 (6), 2121–2168.
- Autor, D.H., Dorn, D., Hanson, G.H., May 2013. The geography of trade and technology shocks in the United States. *Am. Econ. Rev.* 103 (3), 220–225.
- Autor, D.H., Dorn, D., Hanson, G.H., Song, J., Nov 2014. Trade adjustment: worker-level evidence *. *Q. J. Econ.* 129 (4), 1799–1860.
- Baker, S.R., Bloom, N., Davis, S.J., Nov 2016. Measuring economic policy uncertainty*. *Q. J. Econ.* 131 (4), 1593–1636.
- Bernard, A.B., Jensen, J.B., Schott, P.K., Jan 2006. Survival of the best fit: exposure to low-wage countries and the (Uneven) growth of U.S. manufacturing plants. *J. Int. Econ.* 68 (1), 219–237.
- Born, B., Ehrmann, M., Fratzscher, M., 2014. Central Bank communication on financial stability. *Econ. J.* 124 (577), 701–734.
- Caliendo, L., Dvorkin, M., Parro, F., 2019. Trade and labor market dynamics: general equilibrium analysis of the China trade shock. *Econometrica* 87 (3), 741–835.
- Chen, Y., Skiena, S., Jun 2014. Building sentiment lexicons for all major languages. In: *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. Association for Computational Linguistics, Baltimore, Maryland, pp. 383–389.
- Feler, L., Senses, M.Z., Nov 2017. Trade shocks and the provision of local public goods. *Am. Econ. J.: Econ. Policy* 9 (4), 101–143.
- Galle, S., Rodríguez-Clare, A., Yi, M., Jan 2023. Slicing the pie: quantifying the aggregate and distributional effects of trade. *Rev. Econ. Stud.* 90 (1), 331–375.
- Gentzkow, M., Kelly, B., Taddy, M., Sep 2019. Text as data. *J. Econ. Lit.* 57 (3), 535–574.
- Ginsberg, J., Mohebbi, M.H., Patel, R.S., Brammer, L., Smolinski, M.S., Brilliant, L., Feb 2009. Detecting Influenza epidemics using search engine query data. *Nature* 457 (7232), 1012–1014.
- Goldsmith-Pinkham, P., Sorkin, I., Swift, H., Aug 2020. Bartik instruments: What, When, Why, and How. *Am. Econ. Rev.* 110 (8), 2586–2624.
- Hansen, S., McMahon, M., Prat, A., May 2018. Transparency and deliberation within the FOMC: A computational linguistics approach*. *Q. J. Econ.* 133 (2), 801–870.
- Lu, Y., Shao, X., Tao, Z., 2018. Exposure to Chinese imports and media slant: evidence from 147 U.S. local newspapers over 1998–2012. *J. Int. Econ.* 114, 316–330.
- Lucca, D., Trebbi, F., Sep 2009. Measuring Central Bank communication: an automated approach with application to FOMC statements, Technical Report w15367, National Bureau of Economic Research, Cambridge, MA.
- Mikolov, T., Chen, K., Corrado, G., Dean, J., 2013. Efficient estimation of word representations in vector space. In: *International Conference on Learning Representations* arXiv.
- Pierce, J.R., Schott, P.K., Jul 2016. The surprisingly swift decline of us manufacturing employment. *Am. Econ. Rev.* 106 (7), 1632–1662.
- Rehurek, R., Sojka, P., 2011. Gensim–python framework for vector space modelling. *NLP Cent. Fac. Inform. Masaryk Univ. Brno Czech Repub.* 3 (2).
- Saiz, A., Simonsohn, U., Feb 2013. Proxying for unobservable variables with internet document-frequency. *J. Eur. Econ. Assoc.* 11 (1), 137–165.
- Shapiro, A.H., Sudhof, M., Wilson, D.J., Jun 2022. Measuring news sentiment. *J. Econ.* 228 (2), 221–243.
- Tetlock, P.C., Jun 2007. Giving content to investor sentiment: the role of media in the stock market. *J. Finance* 62 (3), 1139–1168.
- Tolbert, C.M., Sizer, M., Tolbert, C.M., Sizer, M., 1996. U.S. Commuting Zones and Labor Market Areas: A 1990 Update, Economic Research Service Staff Paper 9614, United States Department of Agriculture.