

What Triggers Stock Market Jumps?

Scott R. Baker,^a Nicholas Bloom,^b Steven J. Davis^c and Marco Sammon^d

13 April 2021

Abstract: We examine next-day newspaper accounts of large daily jumps in 16 national stock markets to assess their proximate cause, clarity as to cause, and the geographic source of the market-moving news. Our sample of 6,200 market jumps yields several findings. First, policy news – mainly associated with monetary policy and government spending – triggers a greater share of upward than downward jumps in all countries. Second, the policy share of upward jumps is inversely related to stock market performance in the preceding three months. This pattern strengthens in the postwar period. Third, market volatility is much lower after jumps triggered by monetary policy news than after other jumps, unconditionally and conditional on past volatility and other controls. Fourth, greater clarity as to jump reason also foreshadows lower volatility. Clarity in this sense has trended upwards over the past century. Finally, and excluding U.S. jumps, leading newspapers attribute one-third of jumps in their own national stock markets to developments that originate in or relate to the United States. The U.S. role in this regard dwarfs that of Europe and China.

JEL Codes: G12, G17, E58, E62

Keywords: stock market jumps, policy news, market clarity, market volatility, uncertainty, newspapers, human classification

Acknowledgements: We gratefully acknowledge the National Science Foundation, the Sloan Foundation, and Chicago Booth for financial support. We thank Tasaneeya Viratyosin and Jana Obradovic for extensive research support and our team of newspaper coders: Sophie Wang, Alexandra Lee, Naoko Takeda, Abdulla Al-Kuwari, Lucy Duan, Maranna Yoder, TJ Ram, Riley Burton, May Hliang, Jimmy Apffel, Zach Argo, Alex Liang, Kenar Vyas, Meera Desai, Tyler Staggs, Kyle Kost, Ethan Bernheim, Florentin Zander, Alexandra Farmer, Jason Jiang, Robin Gong, Lindsay Roston, Federico Clerici, AJ Qian, Melina Kompella, Heejin Hahn, Devon Shiff, Edina Lee, Sara Holston, Andrea Clerici, Kristina Kolpkova, Alex Zafran, Nikhil Bhatia, Gabe Alon, Garrett Matsuda, Tammy Qiu, Jack Jaffe, Eddie Dinh, Adam Jorring, and Fernanda Kramer.

^a Kellogg School of Management, Northwestern University and NBER

^b Stanford University and NBER

^c Chicago Booth School of Business, Hoover Institution, and NBER

^d Kellogg School of Management, Northwestern University

1. Introduction

What drives big moves in national stock markets? The benchmark view in economics and finance holds that stock price changes reflect rational responses to news about discount rates and corporate earnings. Under this view, we expect big daily moves to be accompanied by readily identifiable developments that affect discount rates and anticipated profitability. Moreover, contemporaneous news accounts should contain information about the proximate drivers of these moves. Of course, stock price behavior may not conform to the benchmark view. Keynes (1936), for example, famously argued that investors price stocks based not on their opinions about fundamental values but on their opinions about what others think about stock values. Even when speculative or irrational forces are in play, however, we expect contemporaneous news accounts to discuss the (perceived) drivers of big market moves. Thus, we turn to newspapers to distill information about what triggers big moves in national stock markets.

Specifically, we examine next-day newspaper accounts of big daily moves (“jumps”) since 1900 in the United States, since 1930 in the United Kingdom, and since the 1980s in 14 other national markets. A threshold of 2.5 percent, up or down, for the U.S. stock market yields 1,150 jumps from 1900 to 2020. These jumps account for only 3.5 percent of trading days but nearly 20 percent of total daily variation (sum of absolute returns) and half of daily quadratic variation (sum of squared returns). Our jump thresholds for other countries range from 2 to 4 percent, with larger thresholds for markets with greater volatility. All told, we examine 6,200 daily stock market jumps across 16 national markets plus another 450 jumps in U.S. bond markets from 1970 to 2020.

Jumps of the size we consider typically attract coverage in leading national newspapers. We locate and read next-day articles about each jump to assess its proximate cause(s), clarity as to cause, and the geographic source of the market-moving news. Our objective is to accurately characterize and code the journalist’s explanation and interpretation of the jump. Trained human readers classify the proximate reason for each jump into 17 categories, one of which is “Unknown & No Explanation.” Readers also code the confidence with which the journalist advances an explanation for the jump and the ease or difficulty of coding the article. For the United States, we examine next-day articles in the *Wall Street Journal*, *New York Times*, *Washington Post*, *Chicago Tribune*, *Los Angeles Times* and other papers. We deploy multiple readers to each paper for each jump so as to

obtain many “reads” per jump.¹ To quantify clarity about the reason for each jump, we combine our data on journalist confidence, ease of coding, pairwise agreement across reads, and whether the journalist advances an explanation for the jump.

Previous studies have also examined news reports to evaluate the drivers of stock market moves. Classic studies by Niederhoffer (1971) and Cutler, Poterba, and Summers (1989) examined major jumps to assess whether they could be explained by identifiable news events, reaching mixed conclusions. Our study advances on earlier work in several respects: scale, encompassing more than 6,000 jumps; scope, spanning 16 national markets and more than 90 years for the U.S. and U.K. markets; detail as to jump reason and geographic origin of market-moving news; and the relationship of jump reason and jump clarity to future market volatility. We also develop a novel approach to quantifying clarity about the forces that drive stock market jumps.

For journalists observing the market in real time, attribution to a clear causal trigger is easy for many jumps and hard for others. To illustrate this point, Figure 1 plots intraday market values at 1-minute intervals on four U.S. jump days. The top panels exhibit large, abrupt intraday moves associated with important pieces of news. In the top left, the market jumped over 3% after the Fed announced a surprise interest rate cut. In the top right, the market plunged 2.5% at the open after an unexpectedly bad employment report. In contrast, the lower panels illustrate two instances that involve large intraday and full-day moves without a clear cause, and for which journalists advanced no explanation. In our U.S. sample, 17 percent of jumps occur for no apparent reason, according to next-day journalistic accounts.

Leveraging our jump-day characterizations, we develop several novel findings. First, upward jumps attributed to policy-related news are more common than downward policy-driven jumps. This pattern holds in every country, and it has strengthened since 1980 in the United States and the United Kingdom, the two countries with pre-1980 coverage. From 1980 to 2020, upward policy jumps are twice as common as downward policy jumps in the United States. Over the same period, downward jumps attributed to non-policy factors are nearly twice as common as upward non-policy jumps. To put the point another way, policy-related developments trigger 43 percent of upward U.S. jumps since 1980 but only 20 percent of the downward jumps.

¹ Given the scale of our data collection efforts, we deploy more than 45 trained human readers comprised of the authors, graduate students, and undergraduate students. Baker et al. (2021) set forth our coding guide and reference manual. Our codings for the *Wall Street Journal* are at <https://stockmarketjumps.com/research/>.

Drilling down, we find that news about monetary policy and government spending is responsible for this pattern. One potential explanation is that positive (negative) monetary policy and government spending surprises are more (less) likely in the wake of bad economic news. Indeed, the share of upward jumps attributed to policy news rises as stock market performance worsens over the preceding three months. And, the share of downward jumps attributed to policy falls as recent stock market performance worsens. This pattern of countercyclical in jump-inducing policy surprises is much stronger after World War II than before.

A skeptic might interpret these patterns as artifacts of how journalists perceive the world and cover the news. Perhaps journalists are prone to credit government policy for upward jumps and to overlook policy mistakes that trigger downward jumps, and perhaps this type of bias manifests itself more powerfully after a period of falling stock prices. While we cannot rule out the possibility of biased coverage, we validate our newspaper-based classifications with reference to FOMC announcements, macroeconomic statistical releases, national election dates, and the industry distribution of jump-day returns. All of our validation exercises support the conclusion that our newspaper-based classifications are informative about the forces that trigger stock market jumps. In addition, our jump classifications have predictive value for future market volatility.

That brings us to our second set of findings: Jumps attributed to monetary policy developments foreshadow considerably less volatility than other jumps. This result holds unconditionally and conditional on controls for the size and direction of jump-day returns and market volatility over the day, week, and month preceding the jump. For example, our conditional forecast of stock market volatility over the next two weeks is two percentage points lower after a jump attributed to Monetary Policy than after other jumps. This gap is three-quarters as large as the time-series standard deviation of average two-week volatility in our sample. We note that monetary policy surprises trigger about 10 percent of all postwar U.S. jumps since 1900 and a similar share of jumps across the other 15 national markets covered by our study.

Our third set of findings pertains to clarity about the reasons for stock market jumps. Our clarity measure fluctuates over time in a positively autocorrelated manner, and it shows a clear upward trend. Over the past 90 years, the share of jumps due to unknown forces fell from about 35 percent to 10 percent in both the United States and the United Kingdom. The other components of our clarity index – journalist confidence, pairwise agreement rates, ease of coding – tell a similar story. As we discuss, there are sound reasons to think the trend toward greater clarity about stock

market behavior reflects a combination of more transparency about corporate performance, better statistical information about the economy, falling communication costs, and the professionalization of news reporting.

We also provide evidence that clarity matters for stock market volatility. In particular, greater clarity about the reasons for a jump today foreshadows lower market volatility over the next week. This result holds in an unconditional sense and conditional on the size and direction of the current jump and market volatility over the preceding day, week and month. Since clarity itself is autocorrelated, this result says that clarity partly accounts for the well-known positive serial correlation of stock market volatility.

Finally, we find that news about U.S. economic and policy developments exerts an extraordinary influence on equity markets around the world. Excluding the United States and focusing on the other 15 national markets covered by our study, news about U.S.-related developments triggers 32 percent of all equity market jumps from 1980 to 2020 (38 percent when dropping jumps due to unknown forces and those for which we find no next-day article.) The U.S. role in this regard dwarfs that of Europe as a whole, even though Europe accounts for a greater share of global output. News about economic and policy developments related to European countries and supranational European institutions seldom drives jumps in non-European countries, with the clear sustained exception of the European sovereign debt crises in the early 2010s. China-related news plays almost no role as a source of jumps in other countries before the mid 1990s, but China-related news has since emerged as an important source of market jumps in other countries.

Our work builds on and contributes to several literatures. There is a wide-ranging literature on how media coverage affects financial markets – contributions include Tetlock (2007), Engelberg and Parsons (2011), and Carlin et al. (2014). Rather than the effects of media coverage on financial markets, we focus on how newspapers interpret stock market jumps. In this respect, we are closer to the classic studies of Niederhoffer (1979), Cutler, Poterba, and Summers (1989) and to more recent work by Manela and Moriera (2017) and Baker et al. (2019), who use newspapers to parse the sources of overall stock market volatility. Relative to these works, we show that different *types* of news events differ in their implications for future volatility, that the informativeness and clarity of newspaper accounts has improved over time, that clarity matters for future market volatility, and that U.S.-related news plays an outsized role in equity markets around the world.

Another literature considers how the clarity of financial writing affects stock returns – prominent works include Li (2008) and Shiller (2017). We contribute to this literature by developing a new approach to measuring clarity about the forces that drive market jumps. Our method is simple and transparent, which facilitates its application to other countries, time periods, and asset markets. We also show that low-clarity jumps are associated with greater post-jump volatility. While we do not pursue it here, our measurement approach opens the door to quantitative studies of how corporate disclosure rules and the accuracy, depth, and timeliness of economic statistics affect clarity about national stock market behavior and thereby influence overall market volatility.

An enormous literature in asset pricing investigates whether, and to what extent, stock market moves can be attributed to news about future cash flows and discount rates. In addition to already-mentioned studies, leading contributions include Shiller (1981) for market-level moves and Roll (1988) for firm-level changes. Many papers study the impact of news releases on market outcomes. Examples include Birz and Lott (2011), Boudoukh et al. (2013), Goldberg and Grisse (2013), Fernandez-Perez et al. (2017), and Fisher et al. (2017)). We proceed in the other direction, starting from large market moves and asking what triggers them according to newspaper accounts.

Finally, much has been written about the outsized role of the U.S. Dollar and Fed monetary policy in the international monetary and financial system. Recent contributions include Obstfeld (2015), Boz et al. (2017), Gopinath and Stein (2018), and Maggiori et al. (2019). These works highlight the role of the U.S. as a global supplier of safe, liquid debt securities, the Dollar's favored status in foreign exchange reserves, the Dollar's prevalence as the currency of denomination in trade invoicing, offshore bank lending and portfolio holdings, and the spillover effects of U.S. monetary policy. We contribute to this literature by showing that news about U.S.-related economic and policy developments plays a uniquely large role as a source of stock market jumps in countries around the world. In this respect, our results reinforce and deepen the results in Ehrmann et al. (2011), who find that spillovers from U.S. bond, equity and money markets to European financial markets are much larger than the other way around.

Section 2 explains how we use newspapers to characterize stock market jumps, describes our jump classification scheme, and undertakes several investigations to assess the quality of our classifications. Section 3 presents several of our main findings about stock market jumps and their relationship to post-jump volatility. Section 4 explains how we measure clarity about the forces that drive particular stock market jumps, documents several empirical properties of our clarity measure,

and provides evidence that greater clarity foreshadows lower stock market volatility. Section 5 applies natural language processing and machine-learning tools to the classification of stock market jumps. The resulting algorithmic approach yields classifications of lower quality and less granularity than the ones produced by our structured human readings. We explain why human readings work better for our classification problem, and we identify circumstances in which the algorithmic approach gets closer to the human-reading approach. Section 6 concludes.

2. Data Creation Process

Before describing our data creation process in detail, we return to the examples in Figure 1. The top left panel of Figure 2 displays an excerpt from an article about the upward jump on 18 April 2001. We classify the reason for this jump as “Monetary Policy and Central Banking,” because the article title and first sentence attribute the jump to a “*surprise rate cut*” by the Federal Reserve. Since the Fed is a U.S. policy institution, the geographic source is the United States. Journalist confidence is “high,” because the article forcefully and unambiguously attributes the jump to the Fed’s decision. Ease of coding is “easy,” because the article is easy to comprehend, the jump reason is easy to discern, and the mapping to our classification of jumps by reason is also easy. The lower right panel of Figure 2 displays an excerpt from an article about the downward jump on 2 July 2009. This article makes clear in the title and first sentence that an “*unexpectedly gloomy jobs report*” triggered the downward jump, which we classify under “Macroeconomic News & Outlook.” The United States is the geographic origin, journalist confidence is “high,” and coding is “easy.”

Figure 3 displays excerpts from two articles about a 5 percent upward jump on 26 December 2018. We classify the jump reason as “Unknown” for the *Wall Street Journal* article, because the reporter writes, “*investors and traders were left scratching their heads to explain the wild swings.*” This passage appears in the third paragraph, reflecting a common practice of placing less-assured explanations further down the article. Since the article says the jump is due to unknown forces, we leave the geographic origin blank. For the *New York Times* article, we classify the primary jump reason as “Macroeconomic News & Outlook” based on “*early reports of a strong holiday-shopping season helped lift the S&P 500 by nearly 5 percent.*” Other papers offer various explanations for this jump. For example, the *Los Angeles Times* offers three explanations over nine paragraphs, including “*a late report that a U.S. government delegation will travel to China.*” Overall, the jump on 26 December 2018 is a low-clarity event in that some papers explicitly attribute it to unknown

forces, and others offer a variety of reasons. Journalists present their explanations for this jump with less assurance and more qualifications. Section 4 below explains how we integrate these aspects of newspaper coverage to quantify clarity about the reason for this jump and others.

2.1 U.S. Stock Market Jumps

To assemble our sample of U.S. jumps, we first identify all days on which the CRSP S&P 500 Value-Weighted Index rose or fell by at least 2.5 percent (close to close). Before 1926, we use the Global Financial Data (GFD) extension of the Dow Jones index. Next, for each jump and newspaper,² we search for next-day articles about the jump, which often post online the prior evening in the internet era. Coders search for articles that mention phrases like ‘stock market,’ ‘wall street,’ ‘S&P,’ or ‘Dow Jones’ in the title, synopsis, or index of descriptive terms. They avoid summaries, abstracts, digests, and the like (articles <300 words). If the search query yields multiple articles for a particular jump and paper, the coder selects the first one unless it proves uninformative. This process yields at least one article for all U.S. jumps since 1926 and almost all before 1926.

Having selected an article, the coder reviews it carefully and – based on the journalist’s characterization – classifies the jump reason into one or more of the following categories (ordered by frequency of U.S. jumps): Macroeconomic News & Outlook, Unknown & No Explanation, Corporate Earnings & Outlook, Sovereign Military & Security Actions, Monetary Policy & Central Banking, Government Spending, Commodities, Regulation, Other Non-Policy reasons, Elections & Political Transitions, Other Policy reasons, Foreign Stock Markets, Taxes, Exchange Rate Policy & Capital Controls, International Trade Policy, Foreign Stock Markets, and Terrorist Attacks and Non-State Violence. We also have a category for No Article Found. The coding guide in Baker et al. (2021) defines each category, includes many examples, explains how to handle jumps attributed to multiple causes, and discusses boundary cases and other challenging cases.

As an example, here is what the coding guide says about how to distinguish ‘Monetary Policy & Central Banking’ from ‘Macroeconomic News & Outlook’:

Some news articles that discuss market reactions to macro developments also discuss the Fed’s normal response to the macro development. Generally, we code an article as Macro News & Outlook if it attributes the market move to news about the macro economy. We code it as Monetary Policy & Central Banking if the article attributes

² For certain exercises, we limit our analysis to results from the Wall Street Journal. This newspaper has the most thorough coverage of financial news and has the most complete archive back to 1900.

the market move to (a) shifts in how the Fed responds to a given macro development or (b) news about unexpected consequences of Fed actions. Take the following two examples:

1. Macroeconomic News & Outlook example: The market moves because it anticipates or speculates (or sees) that the Fed will respond in its usual manner to news about the macro economy. That is, the market anticipates or speculates that the Fed will respond to macro developments according to a Taylor Rule or other well-defined, well-understood description of the Fed's interest-rate setting behavior.
2. Monetary Policy & Central Banking example: The market moves because of a surprise change in the policy interest rate -- i.e., a surprise conditional on the state of the macro economy. From a Taylor Rule perspective, we can think of this change as a new value for the innovation term in the Taylor rule.

We classify the primary reason for each jump into one of our categories and, when warranted by the article's discussion, a secondary reason as well. If an article mentions multiple reasons for a given jump but does not clearly identify the most important one, we treat the order of appearance in the article as a tie breaker. Coders can also designate additional reasons beyond the primary and secondary reasons. Table 1 reports the frequency distribution of primary jump reasons for the United States, the United Kingdom and the "ROTW" (countries other than the United States). We also report the frequency distribution for jumps in the U.S. Treasury bond market. Relative to equity markets, bond market jumps are much more dominated by Macro News & Outlook and news about Monetary Policy & Central Banking.

We also record the geographic source(s) of each jump's primary reason, again based on the explanation offered by the journalist. For instance, we code the United States as the geographic source for a U.S. jump attributed to a Fed policy announcement, while we code the United Kingdom as the geographic source for a U.S. jump attributed to Britain's decision to exit the gold standard. The geographic source can consist of multiple countries or a large region of the world for various reasons that include the outbreak of war between two or more countries, economic developments that simultaneously affect several countries, and policy decisions taken by supranational institutions like the European Central Bank.

We record four additional pieces of information for each article: First, 'Journalist Confidence' is the assurance or certitude with which the article advances an explanation for the jump, which we score on a three-point scale of 1 for low confidence, 2 for medium confidence, and 3 for high confidence. For example, if an article asserts without qualification that bad news about

corporate earnings drove a downward jump, Journalist Confidence is high. Second, ‘Ease of Coding’ reflects the ease or difficulty of discerning and classifying the primary jump reason, also scored on a three-point scale. While Ease of Coding correlates with Journalist Confidence, they are distinct concepts. For example, a journalist may confidently assert an explanation that involves multiple causes that touch on several of our categories. In this case, the primary jump reason may be hard to discern and classify, even though Journalist Confidence is high.³ Third and fourth, the coder paraphrases the journalist’s explanation records the key passage that is the basis for the classification of the primary jump reason.⁴ For example, when the primary jump reason is Taxes, the key passage might say, “The completion of a tax deal between the White House and Congress sent stocks soaring Wednesday.”⁵

For the United States, we conducted a thorough cross-validation with an average of 9.1 coders across multiple newspapers for each day.⁶ Each coder followed the coding procedure outlined above, as detailed in “Coding Large Daily Financial Market Moves - Data Construction Guide”. After all articles were read, we re-examined days where coders disagreed about the primary and secondary cause of the market movement. This happened more often on days that were also coded as having a lower ease of coding and less confidence by the article’s author regarding the driver of the market movement.

To resolve each disagreement, coders re-read the original article and referred to the Data Construction Guide to make sure that the guidelines were being carefully followed. Most disagreements were easily resolved as a coder may have misread an article or misapplied the guidelines from the Data Construction Guide. For articles which still produce disagreement, additional articles in the same newspaper were obtained through the same method as outlined above to seek clarity regarding the primary cause. After these steps were taken, coders still sometimes disagreed. For such days, coders could ‘agree to disagree’ regarding the causes of the stock move and our final dataset reflects such persistent disagreement.

³ The guide contains many examples that illustrate how to score Journalist Confidence and Ease of Coding.

⁴ The third and fourth pieces of recorded information are especially useful for ex post evaluation and refinement of our classification of jump reasons. For example, long after we began our work to classify stock market jumps, we used these fields in Baker et al. (2020) to quickly search for U.S. jumps triggered by pandemics and infectious diseases from 1900 to 2020.

⁵ The third and fourth fields are especially useful for ex post refinements of our jump classifications. For example, long after we began this project, we decided to distinguish

⁶ Appendix Figure A9 summarizes the number of newspapers consulted and codings per jump. Both quantities increase after 1980 when we are able to add the US edition of the Financial Times and the Houston Chronicle.

Finally, before coders started their first coding assignments, they carefully read the coding guide, underwent a half-day training session and then coded 50 WSJ training articles over the next three days. These WSJ training articles had already been coded by us, enabling us to ensure our coders were accurately coding (and to address any issues) before they coded the research sample.

2.2 Foreign Stock Jumps Data

For the US we choose a threshold of a 2.5% daily change in the stock market to define our large jump days to code. This threshold, which covers about 3.5% of trading days from 1900-2020, was chosen to be large enough to ensure the next day newspaper always contained articles discussing the prior days jump. When we extended to other countries, we usually maintained a 2.5% daily return threshold to classify stock market moves as a significant event. For a subset of countries with more volatile stock-markets we increased the threshold, choosing it to cover approximately 2-3% of trading days.⁷

Most foreign countries in our sample only utilized a single paper and about 30-40 years of data. For the UK, however, we conducted a more extensive analysis, with coders searching the Financial Times, The Times of London, The Telegraph and The Guardian the day after any move in the UK stock market larger than 2.5% back to 1930. For the UK, our definition of the aggregate market changed over time: (1) From 1930-1983 we use GFD’s “UK Industrials” index (2) From 1984-1993 we use the percent change in the FTSE 100 index level (3) From 1994-2020 we use the FTSE 100 total return index. Coders searched for the following terms: ‘FTSE’, ‘London stock exchange’, ‘stock market’, ‘equity market’, ‘share prices’. ‘FTSE’ was the most useful keyword in recent years. We mostly use articles longer than 300 words, but for FT articles early in the sample period, shorter articles were more common.

Outside the US and UK coders searched the archive of the newspaper of record for a given country (e.g. the Globe and Mail for Canada). This may take the form of English-language or non-English-language newspaper (e.g. Handelsblatt in Germany or South China Morning Post in Hong Kong). If a non-English-language paper was used, a native speaker of that language was used as a coder. As with the coders for the United States, foreign country coders searched for articles on the day following each jump that mention the stock index in question or the stock market more generally. If the date is a Friday or Saturday, Monday’s paper would be searched, as well.

⁷ Appendix Table A1 lays out the threshold, start date, and primary newspaper utilized for each country.

2.3 Validation of Jump Classifications

A potential concern with our data is the reliability and consistency of our classifications by jump reason. We test for consistency across coders who are investigating a given day's large stock movement by (a) reading articles in the same newspapers and (b) reading articles in different newspapers. Table 2 examines various dimensions of cross-coder ‘agreement’ in categorization. First, we examine the average annual pairwise agreement in primary categorization across all pairs of coders (both within/across newspapers). We find that in the pre-World War II sample, 75% of coders agree on the policy vs. non-policy split, and 43% agree on the 17 granular categories. While this may not seem high, if we randomly assign coders to categories from the unconditional distribution in Table 1, agreement would be only 13%. Based on simulation results, our human coders’ agreement rate is statistically significantly higher than agreement from this random categorization. Further, agreement increased over time, consistent with an increase in the quality of financial journalism. In the post-World War II sample, 80% of coders agree on the policy vs. non-policy split, and there is a 54% agreement rate on the granular categories.

Agreement also increases when considering only pairs of coders reading the same newspaper.⁸ This increase is likely driven by the fact that, for a fraction of the days we study, the cause is ambiguous, leading to be significant differences in how different reporters write about the previous day’s market movements. Suggestive evidence for this is that days in which the articles have lower reported levels of journalist ‘confidence’ also have lower rates of cross newspaper coder agreement. For example, an increase in average reporter confidence of 1 point (on a three-point scale) increases the rate of coder agreement by almost 40%. An increase in the reported ease of coding has an even larger magnitude.

Among papers, agreement is highest for coders of the Wall Street Journal, which we feel has the highest quality financial reporting of all newspapers in our sample. For the WSJ, there is an agreement rate on the policy vs. non-policy split of over 90%, and an agreement rate on the granular categories of almost 80%.

For a subset of categories, we expect that regular information releases drive large stock movements and can use this to test our coding. For instance, we would expect days to be categorized

⁸ The cause can disagree between coders within the same paper if the paper has more than one article on the jump, which may happen in leading financial newspapers (e.g. the Wall Street Journal or Financial Times) on days after major stock-market jumps.

as ‘Elections & Political Transitions’ more often following elections than for the average jump day. Similarly, we would expect a relationship between Federal Reserve announcements and ‘Monetary Policy & Central Banking’ categorizations and high-profile macroeconomic releases (e.g. unemployment numbers and inflation reports) and ‘Macroeconomic News & Outlook’ categorizations.

In Table 3, we demonstrate that these relationships hold statistically, despite coders not directly observing the dates of information releases (i.e. they read only the newspaper article in question and did not separately look up whether the Federal Reserve had made a statement during the previous day). In all cases, the expected categorization is substantially more likely to occur following the public information release. In Appendix A, we extend this validation by examining relative industry returns on jump days, finding results that again support the view that our newspaper-based explanations are informative.

Finally, we compare our classifications of the 50 largest daily U.S. stock market jumps from 1946 to 1987 to those of Cutler, Poterba and Summers (1988), who attribute a “cause” to each one based on *New York Times* coverage. For each of these 50 jumps, we map their “cause” to a primary and, sometimes, secondary category, using our classification scheme. As reported in Appendix Table A3, the resulting CPS primary reason matches our primary reason in 62 percent of cases – 75 percent, when we focus on the subset of 32 jumps with positive values of our clarity measure. For these 32 jumps, there is at least partial overlap between their reasons and ours in 88 percent of cases.

3. Characterizing Stock Market Jumps

3.1 Stock Market Jumps Over Time

Using our human coders, we find a significant amount of variation in the categorical drivers of jumps during the past 120 years. Figure 4 displays the evolution of large daily stock market jumps over time in the United States from 1900 to 2020. Also noted are the fraction of daily jumps that are driven by government policy rather than non-policy causes like news about the economy or corporate earnings, as categorized by coders reading the Wall Street Journal. For a relatively small fraction of articles, the cause of the market’s movement for a given day cannot be determined by coders reading newspaper articles and a categorization of ‘unknown’ is utilized (shaded black).⁹

⁹ For 5 days early in the sample (all pre-1926), we cannot find an article in the Wall Street Journal related to the previous day’s large market movement. This may be due to poor newspaper reporting or could be possibly

In the figure, we see three notable epochs with high jump frequency: the first running from 1929 to the late 1930s, encompassing the Great Depression; the second during and after the Great Recession from 2008-2012; and the third during the COVID-19 pandemic. While the years after the Great Recession (2010-2019) saw only 5.8 jumps per year on average in the United States, there were 36 jumps in 2020. In fact, the month from 24 February to 24 March 2020 includes 18 jumps over 22 trading days, more than any other one-month period in our sample back to 1900. The Panic of 1907, World War I, and the early 1990s are also periods of unusually high jump frequency. Perhaps surprisingly, the Korean War, and the Vietnam War did not produce many large daily stock market jumps in the stock market. During the post-war period, there are long periods with few daily moves large enough to cross our jump threshold.

Figure A10 mirrors Figure 4 using data for stock market jumps in the United Kingdom back to 1930. Here, we find striking differences relative to the United States: The United Kingdom has few jumps in the 1930s but a huge number in the mid 1970s around a severe recession and crisis. The Great Recession and coronavirus pandemic also stand out as periods of high jump frequency.

3.2 Categorical Drivers of Stock Market Jumps

Table 1 displays summary statistics regarding the distribution of the categorical causes of these large stock market movements in the United States in the pre-war and post-war period, the UK from 1930-2020 and from the 1980s onwards in 14 other national markets.¹⁰ This shows that not only has the frequency of large stock market movements fluctuated substantially over time, but the causes of these jumps have changed, as well. For instance, in the pre-1945 period in the United States, agriculture made up a much larger portion of US GDP, so commodities drove a larger share of big stock movements. World Wars I and II contributed to the large number of sovereign military jumps. Finally, the New Deal was responsible for the high share of regulation jumps in the pre-war period. In the postwar period, we see that Monetary Policy was relatively more important, which is consistent with the start of regularly scheduled FOMC meetings in 1981.

driven by measurement error in daily market moves on the part of the DOW-extension pre-1926 when the market was composed of many fewer stocks than today.

¹⁰ Australia, Canada, China (mainland), China (Hong Kong), Germany, Greece, Ireland, Japan, New Zealand, Saudi Arabia, Singapore, South Africa, South Korea and Spain. We have two sets of observations from China, one based on the Hong Kong stock exchange (as covered by the *South China Morning Post*) and one based on the Shanghai stock exchange (as covered by *Shanghai Securities News* and *China Securities Journal*).

We take away two important stylized facts from this table. First, policy news drives a large portion of large daily stock market movements. Over 36% of US jumps are attributed to policy: more than macroeconomic news (23%) and corporate earnings (11%) combined. Globally, 28.6% of jumps are attributed to policy. Second, the distribution of jump causes in the US, the UK and the Rest of the World (ROTW) is relatively similar – in both countries macro news is the largest driver, with corporate earnings and monetary policy also playing a major role. In the ROTW foreign stock markets are the third largest overall mover, reflecting in particular the role of US stock markets in driving global market movements. Third, newspapers fail to advance any explanation for a sizable share of jumps, sometimes explicitly stating that the reasons for the jump are unknown.

The last column of Table 1 provides the categorical attribution for large jumps in US bond markets. These jumps are defined as daily changes in the 10-year constant maturity rate Treasury Bills of more than 15 basis points, although we read and code only a 25% random sample of bond market jumps in 1980-1982. Yields were very high during this period, so there were a huge number of 15 basis-point daily moves. In bonds, we find a significantly different distribution of categories than with equity jumps. Jumps in bond markets are largely driven by macroeconomic news and news about monetary policy, which collectively account for 72 percent of bond market jumps.

3.3 The Geography of News that Triggers National Stock Market Jumps

Looking beyond our categorical classification of jump reasons, we now consider the geography of the news that triggers national stock market jumps. Figure 5 plots timelines of the geographic source of large stock market jumps in the US, showing the predominance of US-related news for US jumps. Sixty-nine percent of US jumps reflect US-related developments, according to next-day accounts in leading U.S. newspapers. The only periods in which the US share persistently falls below 50% are during the First and Second World Wars (when news related to Europe and Asia was especially important), and during the European debt crises in the early 2010s. We present an analogous plot for the UK in Appendix Figure A11. Consistent with the growing global dominance of US financial markets, the UK's share of UK-sourced jumps fell from 76% between the 1930s and the 1960s to around 38% from 2010 onwards. We also see on the top right of Figure 5 that Europe's role in US jumps appears to be falling, while Asia's has recently risen due to the growing influence of China. In the “Other” category, the most common region is the Middle East due to the impact of Gulf wars and OPEC oil shocks on US stock markets.

Figure 6 plots the share of jumps attributed to US-related and Europe-related news back in Australia, Canada, China (HK), China (Shanghai), Japan, New Zealand, Saudi Arabia, Singapore, South Africa, and South Korea. Remarkably, leading newspapers in these countries attribute, on average, 35.3 percent of daily jumps in their own national stock markets to U.S.-related developments. In contrast, they attribute only 5.5 percent of their jumps to Europe-related developments, even though European countries as a whole account for a larger share of global output than the United States. The outsized role of U.S.-related news as a source of daily stock market jumps continues to hold when we include the European countries, as shown in Figure A19.

3.4 Differences between Positive and Negative Stock Market Jumps

Figure 7 reveals the striking fact that policy-driven stock market jumps are disproportionately more likely to be associated with positive stock returns. In the US since 1900, and particularly since 1980, a higher share of positive jumps were due to policy categories than negative jumps. Figure 7 is a binned scatter plot of our policy variable against the daily stock market return. While the slope is positive between 1900 and 1979, the slope becomes much steeper between 1980 and 2020. This suggests that positive jumps are even more likely to be attributed to policy in recent years than in the past. These findings – that policy jumps are more likely to be positive and that this relationship is becoming steeper over time – also hold in the UK (see Figure A12) and for non-US, non-UK countries in our sample (see Figure A13). These patterns are also true looking across jump sizes – since 1980 every size of jump from 2.5% to 3.0%, 3.0% to 4.0%, and above 4% shows a higher share of positive policy jumps than negative jumps (Appendix Table A4).

Examining the individual categories (Appendix Table A5) we see that monetary policy and government spending jumps are the most likely to be positive. In contrast, sovereign military policy tends to be associated with negative stock-market jumps. Since none of these major policy categories has become significantly more positive over time, the rise in the positivity of policy is mostly driven by a changing mix of policy categories. In particular, the two policy categories with the most negative stock market responses – sovereign military action and regulation – decreased in frequency substantially after 1980. And the two most positive major policy categories – monetary and

government spending – have increased in frequency since 1980.¹¹ So, policy has become increasingly positive in the US since 1980 due to rising importance of monetary and fiscal policy and the declining importance of military and regulatory policies as drivers of stock market jumps.

This raises the question as to why monetary and spending shocks are so positively skewed. One explanation is that large monetary and government spending surprises tend to be expansionary in reaction to negative economic news. In particular, major monetary and spending policies are often in response to negative macroeconomic events like the Global Financial Crisis of 2008/09 or the COVID crisis of 2020.

We present evidence for this in Figure 8. The x-axis represents the cumulative return in the CRSP Value-Weighted index over the past quarter (i.e., the 66 trading days before each jump). For the y-axis, we construct a ‘net’ policy score, which is equal to the share of codings on a given day attributed to policy categories if the jump is positive minus the share of codings attributed to policy categories if the jump is negative. Here, the slope is negative and strongly statistically significant (*t*-statistic over 4), providing evidence that policy tends to act countercyclically against the stock market – positive policy jumps are more likely after a quarter of negative returns and negative policy jumps after a quarter of positive returns.

We refine this result in Table 4: in Columns 1-3, we run a regression of positive policy jumps on cumulative returns over the past 66 trading days, and a set of Heterogeneous Autoregressive (HAR) controls (i.e., sum of squared returns over the past day, week, month and quarter). The coefficient on the past quarter returns is negative, but not statistically significant. When we split this regression into two periods, however, it is clear that policy has become more countercyclical: in the post-World War II period, the coefficient becomes negative and statistically significant. This implies that after 1945, there were relatively more positive policy jumps after a series of negative stock returns than in the pre-World War II period.

This increased countercyclicality is also true for contractionary policy. In columns 4-6, we run a regression of negative policy jumps on cumulative returns over the last 66 trading days. We see that in the post-war period the coefficient on past stock market returns has become more positive. This suggests that negative policy is more likely after a series of good returns now than it used to

¹¹ In Table A6, we examine policy jumps by category and by whether the jumps are positive or negative. Notably, in every country in our sample, policy jumps tend to be more positive than negative and this is weakly true for both the Government Spending and Monetary Policy categories.

be. In Appendix Table A7, we show this pattern holds individual for both for monetary and fiscal policy.

One possible explanation for this result is that all jumps (policy and non-policy) have become more countercyclical in nature. To investigate this in columns 7 and 8 we examine non-policy jumps finding that positive non-policy jumps are more likely after past positive returns (the opposite of positive policy jumps), and negative non-policy jumps have no relationship to past returns.¹²

3.5 Differences in Post-Jump Volatility by Jump Reason

The distribution of perceived reasons for stock market jumps differs across countries and varies over time, as we have shown. We now ask whether the contemporaneously perceived reason for a given jump has predictive content for post-jump market volatility. We find that it does. This result continues to hold when we condition on the size and direction of current-day market returns and on market volatility in the days and weeks leading up to the current day.

The main pattern in this regard is that jumps triggered by monetary policy news foreshadow much less volatility than other jumps. To show this pattern, we run regressions of the form,

$$\begin{aligned} \sum_{i=1}^n \frac{r_{t+i}^2}{n} = & \mathbf{a} + \mathbf{b} (\mathbf{r}_t \times \mathbf{1}_{r_t > 0}) + \mathbf{c} (|\mathbf{r}_t| \times \mathbf{1}_{r_t \leq 0}) + \mathbf{d} (\mathbf{r}_{t-1}^2) + \mathbf{e} (\sum_{i=1}^5 \mathbf{r}_{t-i}^2) \\ & + \mathbf{f} \left(\sum_{i=1}^{22} \mathbf{r}_{t-i}^2 \right) + \mathbf{g} (\mathbf{monetary}_t) + \mathbf{h} (\mathbf{allother}_t) + \mathbf{e}_t, \end{aligned}$$

where \mathbf{r}_t is the one-day return on the CRSP value-weighted index, and the dependent variable is the average realized volatility over n days after t . $\mathbf{monetary}_t$ equals the fraction of codings attributed to Monetary Policy on day t , if t is a jump day, and zero otherwise. We define $\mathbf{allother}_t$ analogously for the collection of other jump reasons. We control for positive and negative returns on day t (following Black 1976) and so-called “HAR” variables that let volatility persistence differ by horizon (following Corsi 2009).

Figure 9 plots the estimated \mathbf{g} and \mathbf{h} coefficients and their 95 percent confidence intervals for $n = 1, 2, \dots, 22$ trading days after a jump day. As the figure shows, jumps triggered by monetary policy foreshadow less volatility over the next month than other jumps. At $n = 10$, for example, the conditional forecast of stock market volatility is about 0.75 standard deviations lower after a jump

¹² Note that in columns 7 and 8, we are excluding ‘unknown’ from the non-policy categories (which is why the coefficients are not -1 times the coefficient in the corresponding ‘policy’ columns).

attributed to Monetary Policy than after other jumps. In the appendix, we break out jumps attributed to Macroeconomic News & Outlook (Figure A14), consider a more granular breakdown of other jump reasons (Table A9), distinguish between positive and negative jumps (Figure A15), and split the sample between recessions and expansions (Figure A16). In all cases, we find lower volatility after jumps attributed to Monetary Policy than after other jumps. Thus, the pattern displayed in Figure 9 is robust in several respects.

A natural interpretation of this pattern is that market-moving news about monetary policy often lessens or resolves uncertainty. For example, the outcome of an FOMC meeting may resolve uncertainty about whether the Fed decides to ease or tighten monetary policy. In contrast, it appears that other types of market-moving news are often associated with greater uncertainty. For example, a bad labor market report may trigger a large downward jump in the stock market and raise near-term uncertainty about the economic outlook.

4. Clarity of Stock Market Jumps

4.1 Measurement and Trends in Clarity

Early on in our project we realized there was a wide variation in how clear the cause of some jumps was compared to others. As shown in Figures 1 to 3 some jumps have very clear causes while others are hard for financial journalists to explain. So, we expanded the human analysis of jumps to measure not only the category of a jump, but also the clarity regarding the cause. We create four proxies of clarity and combine these into an overall “Clarity Index”:

- i. Agreement Across Coders: Consider all possible coding pairs for a given jump. (For example, if we have codings by persons 1, 2 and 3, then there are three pairwise codings: (1,2), (1,3) and (2,3). For each pairwise coding, set a measure of agreement $A_{ij}=1$ if i and j agree on the coding, and 0 otherwise. Then compute overall mean pairwise agreement = Sum A_{ij} / N , where the sum is over all i and j for i not equal to j, and N is the number of possible pairwise codings on the data. We expect agreement across newspapers to be lower if the cause of the jump is less clear – each paper may have their own narrative.
- ii. Ease of Coding: When reading the newspaper, each coder reports how easy/difficult it was for them to code the article as a particular cause. On days with a clearly defined cause, we expect both the ease of coding to be high. On other days, the journalist may not clearly list a

particular cause, or put forth a complex cause which coder might have trouble linking to a particular category. On each day, we measure the average ease of coding score.

- iii. Journalist Confidence: When reading the newspaper, each coder reports how confident the journalist was about the cause of the jump. On days with a clear cause, we expect the journalist confidence to be high. On days driven by narratives, the journalist might list several possible explanations or say that the cause of the movement was uncertain. On each day, we measure the average confidence score.
- iv. Share of Unknown Codings: For each coder j , set $Un_j = 1$ if the primary category code is Unknown, zero otherwise. Compute the mean value of Un_j over coders to obtain the Unknown Cause share for the jump. A higher unknown share is less likely tied to concrete news.

Figure 10 plots these four measures over time, showing in all cases a rise in clarity over time (the “share of unknowns” is a reverse clarity measure). We also combine these into a ‘clarity index’. To do so, we normalize each of the four clarity components to have mean zero and standard deviation one, add together these four normalized components, and finally re-normalize the resultant index to have mean zero and standard deviation one. Figure 11 plots this overall clarity index, showing a rise until about 1980 and then an approximately flat index thereafter. This upward trend is not unique to the US, and is mirrored in the UK (see online Appendix Figures A17 and A18).

One plausible reason for an upward drift in the clarity of newspaper articles about stock market jumps are improvements in the quality, scope, and timeliness of statistical information about the U.S. economy. A full account is beyond the scope of our paper, but it is instructive to review developments over time in the Current Employment Statistics (CES) program as an example. Data from the CES and Current Population Survey form the basis for the BLS Monthly Employment Situation Report, a closely watched statistical release about U.S. labor markets that is well known to move stock, bond and currency markets. See, for example, Flannery and Protopapadakis (2002) and Andersen et al. (2007).

The CES program began in 1915 with data from 200 large manufacturing firms.¹³ In its early decades, the program lacked a formal sample design and retained a focus on the manufacturing sector. The BLS began to apply formal sample design methods to the CES around 1950, following a series of memos and testing efforts in the late 1940s. There were significant improvements in CES

¹³ This paragraph draws on Johnson (2016), Kelter (2016) and Mullins (2016).

sample design in 1964 and incremental improvements over the next 25 years. Annual CES benchmarking to universe-level employment data began in 1982, and began moving to a probability-based sample design in 1995, completing the process in 2003. Monthly sample sizes grew from about 107,000 establishments in 1964 to about 620,000 business and government worksites in 2016.

In addition, the growth of the stock market, both in dollar terms and as a fraction of US GDP, has provided additional incentives for understanding and reporting on market-relevant events. For both market participants as well as journalists who cover business and financial markets, greater resources had been made available over time, enabling timelier and more accurate reporting.

One notable contrast in clarity is seen between the two largest financial crises during our sample period. The Great Depression features some of the lowest levels of clarity of jump cause in our sample, while the Great Recession contains some of the highest levels of clarity. Despite both periods possessing extremely high levels of financial market volatility, most of the largest movements during the Great Recession were clearly attributable to a particular cause, while most of the largest movements in the Great Depression lacked a clear attribution to any particular cause in the next day's newspaper reports. Intriguingly, clarity also fell somewhat in 2016 under the Trump administration.

4.2 Validating our Clarity Index

As one validation of the concept of clarity we examine the relationship between the clarity of individual jumps and the concentration of the daily returns within any 5-minute window. The idea builds on Figure 1 that obvious drivers of stock market jumps tend to generate large moves in short time windows.

To do this we regress concentration - the largest movement of the S&P500 over any 15-minute window divided by the total movement during the day – on our clarity index in Table 5. For each day, we calculate the absolute returns in 15-minute intervals, with the first window being 9:30AM to 9:45AM and the final window 3:45PM – 4:00PM. We then divide the largest absolute move by the sum of all the absolute moves to obtain our concentration measure, with 1 being the highest value (the entire days move happened in one 15-minute window) and 1/26 the lowest value (equally spaced movement across each 15-minute window).

In column (1), without any controls, we see that concentration is highly significantly related to clarity with a t-statistic over 4. Given the mean value of concentration of 0.150 and a standard

deviation of 0.056, this result implies that a two-standard deviation shock to clarity is associated with an increase in concentration of 0.0436, or a 0.78 standard deviation increase. In column (2) we add a full set of controls for returns, absolute returns and prior volatility and find the results are similar. In columns (3) to (6) we examine each individual component of the clarity index and find the expected coefficient.

Table 5 shows that days with a single sudden burst of trading in a single direction tend to be the most ‘clear’ (e.g. the top two days in Figure 1), while days that vacillate back and forth or drift throughout the day tend to have lower clarity according to our approach (e.g. the lower two days in Figure 1). Moreover, as we demonstrate in the following section, these differences in stock market behavior are correlated with clarity not only on the day of a given large stock market jump, but are persistently different for weeks after.

4.3 Jump Clarity and Volatility

In Figure 12, we compare absolute returns around high and low clarity jumps, defined as days with above/below median clarity for the particular time period studied. In the all-years sample, we find that the mean absolute return is significantly higher both in a +/- 3-day and a +/- 22-day window around low clarity jumps than high clarity jumps. There appears to be a significant forward and backward relationship between lower clarity and higher stock-market volatility. That is, jumps that are harder to explain are both proceeded by and followed after by significantly higher stock-market volatility. This is also shown in Figure 13 where we see clarity itself is also persistent, suggesting the markets experience bouts of lower and higher clarity, moving alongside periods of lower and higher volatility.¹⁴ Indeed, this suggests that one reason for the persistence of volatility in financial markets may be the persistence of clarity of the factors driving markets.

Many high-clarity jump days are driven by news about government policies. Figure 14 plots the distribution of clarity for selected policy categories against all non-policy categories, excluding unknown and no article found. Clarity is significantly higher for jumps attributed to sovereign military action (about 1 standard deviation of clarity), monetary policy and government spending (about half a standard deviation of clarity) than all the non-policy categories. So, while many recent studies investigate the role of policy in driving economic uncertainty (e.g., Baker et al. 2016), our

¹⁴ Appendix Table A8 includes more controls to account for the time series trend in clarity, jump categories and the time between jumps, showing the persistence of clarity is robust to including all these controls.

results suggest that policy-driven jumps tend to involve greater clarity and induce less future volatility than other jumps.

In Table 6, we regress future stock market volatility on our clarity index. The first measure of volatility we use is the average squared returns over the five days following the jump. We see a highly significant result in column (1), the specification without any controls. A two standard deviation drop in clarity is associated with an increase in volatility of 10, or a 0.625 standard deviation increase. This suggests that after days in when the movement in the stock-market was hard for journalists to explain there is greater subsequent stock-market volatility. One natural interpretation is that lower clarity events are more difficult for the market to parse, leading to greater future volatility. This is consistent with Carlin, Longstaff and Matoba (2014) who find that increases in disagreement predict future realized volatility. Indeed, clarity and disagreement are likely related, noting in particular our clarity measure is based in part of the extent of agreement within and across newspapers. This result is robust to controlling for the size of the jump, in column (2), and adding HAR controls in column (3).

Our second measure of volatility is the average cross-sectional standard deviation of individual stock returns over the five days following the jump. Specifically, this is the value-weighted standard deviation of returns across all ordinary common shares in CRSP in a given day. Consistent with the first three columns, cross-sectional standard deviation is lower after higher clarity jumps. This suggests that higher clarity events tend to have a similar effect on all firms, while there are more winners and losers after low clarity events. In unreported results, we find that future aggregate volatility and cross-sectional standard deviation are negatively related to each of our individual clarity index components as well.

5. Algorithmic Jump Classification

Overall, we have found that assessing the causes and clarity of large stock market jumps can yield insights into both long-run macroeconomic trends and future stock market volatility around the world. However, given the major costs and time involved with running large-scale human evaluations in order to accurately code thousands of individual daily stock market movements, we are only able to analyze a small fraction of daily market movements in a small fraction of countries. Thus, it is natural to attempt to approach the question using automated textual-analysis tools which would allow for a much greater scale.

5.1 Barriers to Algorithmic Jump Classification

There are a number of reasons to be wary of an automated approach to jump day classification, at least when starting with the blank slate of a database of newspapers and stock market movements.

For instance, using no other input, Latent Dirichlet Allocation (LDA) (see Blei et al. (2003)) can separate newspaper articles into N distinct topics composed of different weights on different sets of terms, but these may not be able to be mapped to categories that humans may find useful or applicable to research. For instance, researchers may be interested in understanding how trade policy drives stock market movements, but a computer may not isolate this category as a distinct factor given the small number of large stock movements driven by trade policy.

This problem is compounded when focusing on large stock market movements. Such a restriction reduces sample sizes and makes any automated approach more prone to issues of overfitting. In addition, the substantial divergence in frequency that each category appears can cause issues with the loss-functions of many off-the-shelf techniques, pushing algorithms towards a tendency of defaulting to the most common categories.¹⁵

These issues are amplified by the fact that language employed by journalists and members of the financial industry have changed significantly over time. The choice of words that describe a large stock move caused by ‘Corporate Earnings’ or ‘Trade & Exchange Rate Policy’ vary widely from 1910 to 2010. This is due both to changes in common terminology over time but also to the fact that the institutional framework of business, government, and financial markets has changed substantially.¹⁶ The composition of journal articles has also changed significantly over the past 120 years, with recent articles being more focused on a particular aspect of the market while earlier newspapers often touched on all facets of both financial and industrial markets in a single article.

¹⁵ One may attempt to gain granularity by increasing the number of dimensions to attempt to fit over (eg. moving from single words to 2-grams or n-grams in order to separate ‘war’ from ‘trade war’ or ‘deficits’ from ‘trade deficits’), but this decreases the generalizability of the resultant classification system out-of-sample. While the automated system may perform well when automating the bifurcation of stock moves into two types of explanations, attempting to split the data into 10-20 categories that exhibit hugely different base rates tend to produce substantial Type 1 and Type 2 errors.

¹⁶ These changes span the creation of the Federal Reserve, the creation and end of countries, the end of the gold standard, the rise and fall of industries, and the broad innovations in financial reporting requirements and new trade agreements.

Automated categorization is also limited to the quality of the PDF files being converted to text. Earlier years (e.g. pre-1940), in particular, suffer from poor image quality which results in less-than-perfect translation into machine-readable text.

5.2 TF-IDF Categorization Using Human-Coded Training Sample

Here we develop a first attempt at categorizing jump day articles in an automated fashion. We work to ‘rank’ the most likely categories for each day based on the raw text of the newspaper articles that were used by our human coders utilizing a TF IDF approach.¹⁷ The full details of this approach are laid out in the Appendix.

For each day, this approach allows us to rank the probability of each category being the correct category in an out-of-sample test. The category with the highest sum will be given rank 1, second highest rank 2, etc. Across our entire sample, our average ranking of the true category is 2.5. That is, on average, the category our human coders identify as the correct cause is often only ranked as the second or third most likely category based on the algorithmic reading.

As mentioned previously, one large concern has to do with the evolution of language and journalistic practices over time. We split our sample into four periods, each containing one fourth of the total jump days in the United States since 1900: 1900-1931, 1932-1939, 1940-2007 and 2008-2020. We repeat our ranking classifier on each sub-sample using a leave-one-out methodology for out-of-sample categorizations. While the oldest sub-sample tend to see an average ranking of approximately 3, the most recent sub-sample has an average ranking of approximately 1.5 (relative to a best-possible ranking of 1). This reflects the tendency for more recent articles to be written in a clearer and more focused fashion, allowing for greater differentiation between articles in terms of the cause for the day’s stock move. This tendency mirrors the evolution of our other measures of human-coded ‘clarity’ over time, showing that automated classification reveals a similar increase.

While this was only a cursory evaluation of text-to-data methods for evaluating news articles we hope our set of almost 1,200 US and 4,000 Rest of World coded alongside the PDFs of the underlying articles will provide a corpus of text to develop more sophisticated methods by other researchers in future.

¹⁷ Here, we restrict our analysis to the Wall Street Journal, for which we can access the raw text of each article back to 1900.

6. Conclusion

We examine newspapers the day after major stock-market jumps to catalog the proximate cause, clarity as to cause, and geographic source of the market-moving news since 1900 in the United States, 1930 in the United Kingdom, and the 1980s in 14 other national markets. We find four main results. First, policy causes a significantly higher share of positive than negative stock market jumps. This result holds in all 16 national markets we cover, and it becomes stronger after World War II in the United States and the United Kingdom. Monetary policy and government spending jumps are the most strongly over-represented in positive jumps, suggesting major policy announcements are usually in response to negative economic news. We also provide direct evidence that policy-induced jumps have become more countercyclical over time – and that this effect is concentrated in jumps attributed to monetary and fiscal policy.

Second, jumps caused by non-policy events lead to higher future stock-volatility, while jumps caused by policy events, monetary policy in particular, lead to lower volatility. This suggests while monetary policy surprises lead to stock-market jumps, they may reduce future volatility. Third, the clarity of the cause of stock market jumps rises substantially over the past 90 years in the United States and the United Kingdom, because information sources have deepened and financial markets have become more transparent. Greater clarity is associated with lower levels of post-jump volatility, suggesting that trend increases in clarity have been a force for less market volatility.

Finally, we find the United States exerts an extraordinary influence on national equity markets around the world. From 1980 to 2020, 32 percent of all jumps in non-U.S. stock markets were triggered by news emanating from or about the United States. This assessment reflects the reportage in leading own-country newspapers about their national stock markets. The U.S. role in this regard dwarfs that of Europe and China. Jumps in other countries attributed to China-related developments were rare before the mid 1990s but have become much more frequent in recent years. This aspect of our findings highlights another facet of the uniquely large role played by U.S. economic and policy developments in the international monetary and financial system.

References

- Andersen, T.G., Bollerslev, T., Diebold, F.X., Vega, C., 2007. Real-time price discovery in global stock, bond and foreign exchange markets. *Journal of International Economics* 73, 251–277.
- Baker, S.B., Bloom, N., Davis, S.J., Sammon, M., 2021. Coding large daily financial market moves: Data construction guide. URL: <https://stockmarketjumps.com/files/newguide.pdf>.
- Baker, S.R., Bloom, N., Davis, S.J., Kost, K., 2019. Policy News and Stock Market Volatility. NBER Working Paper No. w25720.
- Baker, S.R., Bloom, N., Davis, S.J., Kost, K., Sammon, M., Virayosin, T., 2020. The unprecedented stock market reaction to COVID-19. *The Review of Asset Pricing Studies* 10, 742–758.
- Barro, R.J., 1990. The stock market and investment. *The Review of Financial Studies* 3, 115–131.
- Birz, G., Lott, J.R., 2011. The effect of macroeconomic news on stock returns: New evidence from newspaper coverage. *Journal of Banking and Finance* 35, 2791–2800.
- Black, F., 1976. Studies of stock market volatility changes, in: Proceedings of the American Statistical Association Business and Economic Statistics Section, pp. 177–181.
- Blei, D.M., Ng, A.Y., Jordan, M.I., 2003. Latent Dirichlet allocation. *The Journal of Machine Learning Research* 3, 993–1022.
- Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31, 307–327.
- Bollerslev, T., Engle, R.F., Nelson, D.B., 1994. ARCH models. *Handbook of Econometrics* 4, 2959–3038.
- Boudoukh, J., Feldman, R., Kogan, S., Richardson, M., 2013. Which news moves stock prices? A textual analysis. NBER Working Paper No. w18725.
- Boz, E., Gopinath, G., Plagborg-Møller, M., 2017. Global trade and the dollar. NBER Working Paper No. w23988.
- Brogaard, J., Detzel, A., 2015. The asset-pricing implications of government economic policy uncertainty. *Management Science* 61, 3–18.
- Campbell, J.Y., Lettau, M., Malkiel, B.G., Xu, Y., 2001. Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *The Journal of Finance* 56, 1–43.

- Carlin, B.I., Longstaff, F.A., Matoba, K., 2014. Disagreement and asset prices. *Journal of Financial Economics* 114, 226–238.
- Corsi, F., 2009. A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics* 7, 174–196.
- Croux, C., Reusens, P., 2013. Do stock prices contain predictive power for the future economic activity? A Granger causality analysis in the frequency domain. *Journal of Macroeconomics* 35, 93–103.
- Cutler, D.M., Poterba, J.M., Summers, L.H., 1988. What moves stock prices? NBER Working Paper No. w2538.
- Dougal, C., Engelberg, J., Garcia, D., Parsons, C.A., 2012. Journalists and the stock market. *The Review of Financial Studies* 25, 639–679.
- Ehrmann, M., Fratzscher, M., Rigobon, R., 2011. Stocks, bonds, money markets and exchange rates: measuring international financial transmission. *Journal of Applied Econometrics* 26, 948–974.
- Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50, 987–1007.
- Fama, E.F., 1981. Stock returns, real activity, inflation, and money. *The American Economic Review* 71, 545–565.
- Fernandez-Perez, A., Frijns, B., Tourani-Rad, A., 2017. When no news is good news - The decrease in investor fear after the FOMC announcement. *Journal of Empirical Finance* 41, 187–199.
- Fischer, S., Merton, R.C., 1984. Macroeconomics and finance: The role of the stock market. NBER Working Paper No. w1291.
- Fisher, A.J., Martineau, C., Sheng, J., 2017. Media attention, macroeconomic fundamentals, and the stock market. University of British Columbia Working Paper.
- Flannery, M.J., Protopapadakis, A.A., 2002. Macroeconomic factors do influence aggregate stock returns. *The Review of Financial Studies* 15, 751–782.
- Glosten, L.R., Jagannathan, R., Runkle, D.E., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance* 48, 1779–1801.

- Goldberg, L.S., Grisse, C., 2013. Time variation in asset price responses to macro announcements. NBER Working Paper No. 19523.
- Gopinath, G., Stein, J.C., 2018. Banking, Trade, and the making of a Dominant Currency. NBER Working Paper No. w24485.
- Guo, H., 2002. Stock market returns, volatility, and future output. *Federal Reserve Bank of St. Louis Review*. 84, 75–85.
- Gürkaynak, R.S., Sack, B.P., Swanson, E.T., 2005. Do actions speak louder than words? The response of asset prices to monetary policy actions and statements, in: *International Journal of Central Banking*. volume 1.
- Johnson, N., 2016. One hundred years of Current Employment Statistics data collection. *Monthly Labor Review* 139, 1.
- Kelly, B., Pástor, L'., Veronesi, P., 2016. The price of political uncertainty: Theory and evidence from the option market. *The Journal of Finance* 71, 2417–2480.
- Kelter, L.A., 2016. One hundred years of Current Employment Statistics—the history of CES sample design. *Monthly Labor Review*.
- Keynes, J.M., 1936. The general theory of employment, interest, and money. Palgrave Macmillan, London.
- Li, F., 2008. Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics* 45, 221–247.
- Lucca, D.O., Moench, E., 2015. The pre-FOMC announcement drift. *The Journal of Finance* 70, 329–371.
- Maggiori, M., Neiman, B., Schreger, J., 2020. International Currencies and Capital Allocation. *Journal of Political Economy* 128, 2019–66.
- Manela, A., Moreira, A., 2017. News implied volatility and disaster concerns. *Journal of Financial Economics* 123, 137–162.
- Moreira, A., Muir, T., 2017. Volatility-managed portfolios. *The Journal of Finance* 72, 1611–1644.
- Mullins, J.P., 2016. One hundred years of Current Employment Statistics—an overview of survey advancements. *Monthly Labor Review* 139, 1.
- Niederhoffer, V., 1971. The analysis of world events and stock prices. *The Journal of Business* 44, 193–219.

- Obstfeld, M., 2015. Trilemmas and Trade-Offs: Living with Financial Globalisation. BIS Working Papers No. 480.
- Pástor, L'., Veronesi, P., 2012. Uncertainty about government policy and stock prices. *The Journal of Finance* 67, 1219–1264.
- Pástor, L'., Veronesi, P., 2013. Political uncertainty and risk premia. *Journal of Financial Economics* 110, 520–545.
- Roll, R., 1988. R2. *The Journal of Finance* 43, 541–566.
- Russell, S., Norvig, P., 2002. *Artificial Intelligence: A Modern Approach*. 2nd ed., Prentice Hall, Upper Saddle River, NJ.
- Shiller, R., 1981. Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends? *American Economic Review* 71, 421–436.
- Shiller, R.J., 2017. Narrative economics. *American Economic Review* 107, 967–1004.
- Tetlock, P.C., 2007. Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance* 62, 1139–1168.

Appendix

A1. Industry-Level Returns Validation

For some daily stock market jumps, the explanation offered in next-day newspaper accounts implies an amplified or dampened response of equity returns in particular industries to the news that moved the overall market. Consider two examples, the first involving bank stocks and the second involving defense-industry stocks

- Example 1, Banks: During the GFC, the stock market responded positively to upward revisions in the likelihood or generosity of bank bailouts. For this type of jump, we expect an even more favorable response for Bank stocks. That is, the response of Banks is **amplified** relative to the overall market response.
- Example 2, Guns: When bad news about the likelihood or duration of the Iraq war generated a negative jump, we expect the response for Guns (defense firms) to be **dampened** relative to the overall market response. While a longer war may be bad for the overall U.S. economy, it is less bad (or even good) for Guns.

These examples suggest that we can test whether newspaper-based explanations are accurate by examining whether their implications for relative industry-level returns hold in the data.

To do so, we proceed as follows. First, we work with the daily industry-level equity returns data constructed by Gene Fama and Ken French, which are available at https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. Let R_{it} = the daily return for industry portfolio i on day t as measured by Fama and French.

Second, we use the detailed explanations offered in next-day newspaper accounts – as recorded by our human coders – to identify instances in which particular industries should have an amplified or dampened return response if the newspaper explanation is accurate. Using these detailed explanations, we construct an industry-level variable Tri_{it} that takes on three possible values for each industry i on each jump date t , as follows:

$$\begin{aligned} Tri_{it} &= 1, \text{ if the detailed description for } t \text{ implies an amplified response of } R_{it}; \\ &= -1, \text{ if the detailed description for } t \text{ implies a dampened response of } R_{it}; \\ &= 0, \text{ otherwise.} \end{aligned}$$

In constructing this variable, we take a conservative approach: We set Tri to 1 or -1 based on the Primary jump reason only. We set Tri to 0 when the detailed explanation for the jump involves an overly broad industry group. For example, “Manufacturing” maps to at least 15 of the 49 industry groups and is too broad for our purposes.¹⁸

Most jump-day explanations do not map readily to a particular industry. Sometimes, we assign 2 industries to a given jump. Most, but not all, of these dual assignments involve Sovereign Military Jumps, which implicate both Guns and Aerospace. Among our 339 jumps from 1960 to 2016, we obtain 115 Jump-by-Industry observations with nonzero Tri values, as follows: 38 nonzero values for Banks, 19 for Guns, and 16 for Aerospace. Several other industries had fewer than 10 nonzero Tri values: Oil, Coal, Building Materials, Construction, Autos, Chips, Hardware, Household Goods, Software, and Electrical Equipment.

Third, we test whether the implications of newspaper accuracy for relative industry-level returns hold in the data. In our one-industry-at-time approach, we fit the following regression model by OLS to daily returns data for a given industry i ,

$$R_{it} = \alpha + \beta MR_t + \delta Tri_{it} + \gamma Tri_{it} MR_t + \epsilon_t,$$

where MR_t = the daily return on the overall market portfolio on day t . The chief coefficient of interest is γ , which tells us whether the relative industry- i return is amplified or dampened on particular jump days. The null hypothesis is $\gamma = 0$. Newspaper accuracy implies the alternative hypothesis, $\gamma > 0$. The specification includes a control for the market return, because industry i may be relatively sensitive or insensitive to market returns for reasons apart from the ones identified in our newspaper explanations on jump days.

We report the estimated γ coefficient in this regression for the Banks industry in columns (1) and (2) of Table A2. We soundly reject the null hypothesis in favor of the alternative, as seen by the positive sign and statistical significance of the γ coefficient. The estimated value for γ in Column (1), for example, says the return for Banks is amplified by 80 percent relative to the average market return on jump days with $Tri_{Banks} = 1$. Thus, the results in Columns (1) and (2) strongly support the view that next-day newspaper explanations are accurate as to the reason for the jump – at least for those jump explanations that imply an amplified response for Banks.

¹⁸ In practice, Tri typically takes on only two values (0 and 1, or 0 and -1) for a given industry. However, when pooling over industries to get additional power in the regression test below, we will need the trichotomous variable.

As it turns out, Banks is the only industry with a large enough number of non-zero Tri values to yield reasonably precise estimates of γ . Thus, we also fit a multi-industry regression specification, as follows:

$$R_{it} = \sum_i \alpha_i + \sum_i \beta_i MR_t + \sum_i \delta_i Tri_{it} + \gamma Tri_{it} MR_t + \epsilon_t.$$

When fitting this regression, we pool over all industries with at least one nonzero Tri value.

Columns (3) and (4) in Table A2 report the results. Again, we soundly reject the null hypothesis in favor of the alternative, and the estimated value for γ implies a large amplification/dampening effect on returns in those industries that, according to the newspaper-based explanation, should experience an amplification/dampening effect.

In summary, the results in Table A2 provide evidence that next-day newspaper accounts contain meaningful explanations for large daily moves in national stock markets. This evidence about industry-level returns on jump days complements the evidence in Table 3 discussed in the main text. In particular, we stress that Table 3 and Table A2 provide two distinct types of evidence that validate our newspaper-based classifications for jump reasons, and the newspaper explanations themselves.

A2. Algorithmic Stock Market Jump Categorization

We start by OCRing the full text of each Wall Street Journal (WSJ) article. Unlike our other newspapers, we only have 1 WSJ article per day, as part of an experiment to explicitly measure differences among coders reading the same articles in the same paper, rather than reading different articles from the same paper. For most supervised machine learning algorithms, we would like to have exactly one category per day in the training sample. For days where the WSJ coders agreed, this is straightforward. If they disagree, however, we take the category with the highest average score among categories, if the highest average score is above a certain threshold. In this subsection, we make that threshold 0.5, so at least one coder must assign it a lone primary and the other must assign it at least as a secondary category. If no category on a given day crosses this threshold, that day is dropped from the sample.

We then clean the articles by removing all (1) non-English words, which are usually OCR errors from early in the sample when the scanned articles are of lower quality (2) words that are parts of headers/footers generated by ProQuest when the articles are saved as PDFs (3) stop words using the NLTK toolbox in Python. We then do additional cleaning based on the algorithm in Loughran

and McDonald (see <https://sraf.nd.edu/textual-analysis/resources/> for detailed notes on their cleaning procedure) to make the punctuation meaningful, making it easier to break the document into sentences. Finally, we use the Porter Stemmer to reduce all words to their stems.

After cleaning the articles, we extract the first 200 words of each article. This has two main benefits: (1) It makes all the articles the same length, which is useful when doing tf-idf to prevent biases caused by differences in document length and (2) many articles, especially early in the sample, discuss several topics, including those unrelated to the jump. The first 200 words are usually the most relevant for categorizing the article. Finally, we require that words appear in a category at least 3 times, and overall at least 5 times.

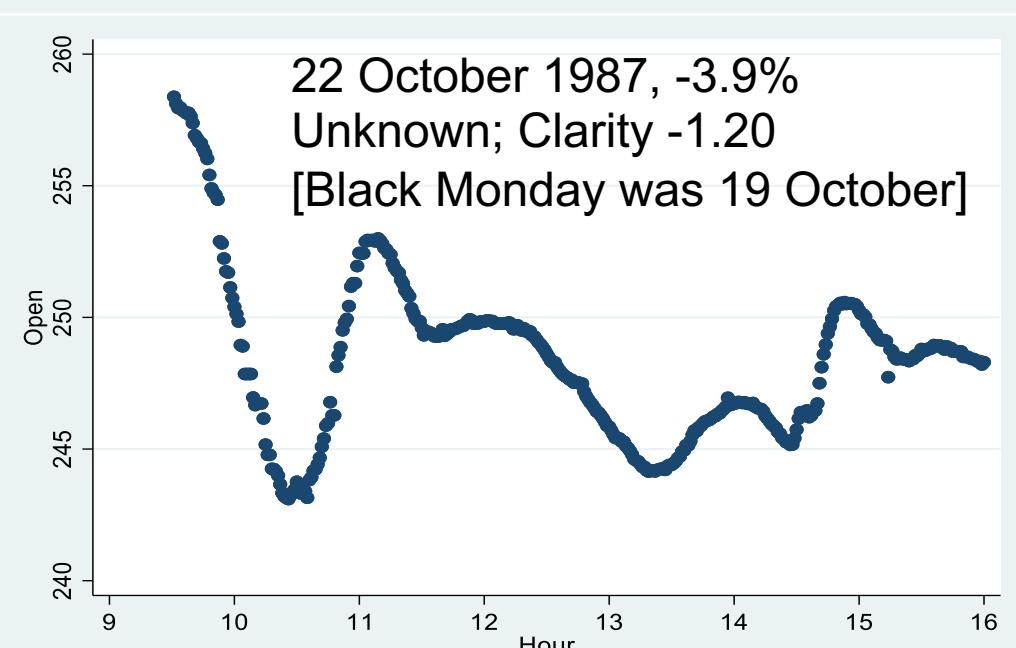
Having cleaned the text data, we compute a tf-idf score for each word in each document. tf is computed within an article, while idf is computed across all articles that survive the filters described above. We then use these scores to perform a ‘leave-out-one’ classification of each article. To do this, we take the entire corpus, excluding the article we are trying to classify. We then take all the unique words in those articles, and sort on the average tf-idf score for these words across articles in each human-classified category. Finally, we take the top 100 words for each category from this sorting: these are the words we associate with each category. For example, for Commodities the top word is ‘wheat’, while for Sovereign Military Action the top word is ‘germani’ (stem of Germany).

As an alternative to the TF-IDF approach, we use our Wall Street Journal codings as the training sample for a Naïve Bayes Classifier (see, for example, Russell and Norvig (2003)). To reduce overfitting, we follow the same procedure described above when constructing the category ranking. The main filters include removing stop words, words that appear in fewer than 5 articles, and words that appear in more than 70% of articles (ie. those with low signal-to-noise ratio). In-sample, the algorithm can fit nearly 100% of articles, but allowing this amount of flexibility may drive overfitting issues. To test for over-fitting, we measure the out of sample performance of the model. For each day, we fit the Bayes Classifier on all other days and then pass that day’s article into the classifier. To account for differences in base rates across categories, we restrict classification among those categories with a sufficiently large sample and similar base rates: Corporate Profits, Government Spending, Macroeconomic News & Outlook, Monetary Policy and Sovereign Military Actions. Although there are a significant number of jumps classified as Unknown, we omit this category, as it adds a noise to out of sample classifications. With this approach, we fit 63% of articles. On average, the Bayes Classifier works better out of sample than randomly picking categories from

the unconditional distribution (which would achieve a match rate of 31%), but the fit is far from perfect.

Restricting to the post-1984 to isolate the sample in which we can obtain the text directly rather than OCRing PDF files, the fit improves slightly. However, this reveals a significant problem: because many of the categories are sparse, the model almost always guesses the modal category of ‘Macroeconomic News & Outlook’. As discussed above, while it is possible to improve the out of sample fit by stemming words and trying to identify ‘relevant’ pieces of long articles (especially in the pre-World War II period), there is a limit to how good the out of sample fit can be with the ‘bag-of-words’ approach.

Figure 1: Intra-Day Moves Often, But Not Always, Point to the Likely Jump Reason



Notes: Each panel plots the S&P 500 index at 1-minute intervals from market open to close on the indicated date. We also report the percent change from the previous-day close to the current-day close, the primary jump reason (as classified by our human readers), and our measure of clarity as to jump reason. The clarity measure is standardized to mean zero and unit standard deviation. The top two panels also report the specific event that, according to newspaper accounts, triggered the jump.

Figure 2: Two Examples of Newspaper Articles about High-Clarity Jumps

Stock Prices Soar, as Investors Embrace a Surprise Rate Cut

By E.S. Browning Staff Reporter of The Wall Street Journal

Updated April 19, 2001 4:22 am ET

SAVE PRINT TEXT

Another surprise interest-rate cut by the Federal Reserve sparked another strong rally in the stock market, with the Nasdaq Composite Index surging 8.1% and the Dow Jones Industrial Average rising nearly 4%.

The question for many investors: Is this rally for real, in contrast with several other short-lived run-ups since stocks began their bear-market drop last year?

The answer, traders and investors say, may depend on whether investors are more fearful of missing out if the market keeps going up, or more worried that the economic outlook will remain cloudy.

Bulls have been encouraged to see stock prices "reacting extremely well compared to the earnings numbers we are seeing," said Tim Heekin, director of trading at San Francisco investment bank Thomas Weisel Partners. Skeptics, however, say they are stunned by the idea that investors would jump back into tech stocks, in particular, after their collapse of the past year.

For the WSJ article at right, we code the primary jump reason as **Macro News and Outlook (Non-Policy)**, because the drop is clearly linked to the poor jobs report. Geographic source is the **US**. Journalist confidence is high, and ease of coding is **Easy**.

For the WSJ article at left, we classify the primary jump reason under **Monetary Policy & Central Banking (Policy)**, because the article links the rise to the Fed's surprise interest rate cut. Geographic source is the **United States**, because the Fed is a U.S. policy-making institution. Journalist confidence is **High**, as the article explicitly links the move to the rate cut. Ease of coding is **Easy**.

Dow Drops 223.32 and Oil Slides --- Many Investors Sell Stocks, and Flock to Treasurys, After Soft Jobs Report

Lobb, Annelena; McKay, Peter A Wall Street Journal, Eastern edition; New York, N.Y. [New York, N.Y.]03 July 2009: C.1.
THE WALL STREET JOURNAL.

[Full text](#) [Abstract/Details](#)

An unexpectedly gloomy jobs report heightened anxiety that the economy mightn't be recovering as well as expected, prompting investors to sell stocks and commodities and flee to haven investments.

The Dow Jones Industrial Average slid 2.6%, the biggest decline ahead of a July 4 holiday in at least 50 years. The Dow closed at 8280.74, down 223.32 points, its lowest close since May 22 and the third consecutive week of declines. The New York Stock Exchange extended trading for 15 minutes at the end of the day because of a computer glitch.

Investors also abandoned commodities, reflecting the diminished optimism for economic growth and demand for raw materials. Crude slumped \$2.58, or 3.7%, to \$66.73 a barrel.

Instead, investors sought the relative safety of U.S. Treasurys and the U.S. dollar. The benchmark 10-year Treasury added 14/32 to 96 30/32, pushing down the yield to 3.494%. The dollar gained 1% against the euro and changed hands at 1.40 per euro late Thursday.

The 467,000 jobs lost in June surprised investors and fueled worries about the strength of the economy. After soaring from a low reached on March 9, stocks had plateaued. The jobs report came on the eve of earnings season, which begins next week with the report of Alcoa. Analysts have begun to worry that, even with the recent decline, stock investors may be overly optimistic about a second-half recovery.

Figure 3: Two Examples of Articles (from Different Papers) about a Low-Clarity Jump

For the WSJ article below, we code the jump reason as **Unknown**, because “traders and investors were left scratching their heads.”

U.S. MARKETS

Dow Industrials Leap More Than 1,000 Points

Rebound pulls Dow industrials, S&P 500 from brink of bear market

By Jessica Menton
Updated Dec. 26, 2018 11:07 p.m. ET

The Dow Jones Industrial Average surged more than 1,000 points for the first time in a single session Wednesday, rebounding after a bruising four-day selloff put the blue-chip index and the S&P 500 on the brink of a bear market.

All 30 stocks in the Dow industrials notched gains, as did each of the 11 sectors in the broader S&P. Shares of Amazon.com, Facebook and Netflix climbed more than 8%, while retailers including Kohl's and Macy's rallied as early data on the crucial holiday shopping season appeared robust. Energy stocks including Exxon Mobil and Chevron, meanwhile, rose alongside a nearly 9% climb in oil prices.

But as in many of the volatile days that have characterized markets since the end of September, investors and traders were left scratching their heads to explain the wild swing, with the Dow adding nearly 450 points in the last hour of the session.

The New York Times

Stocks Bounce Back From Edge of Bear Market

S&P 500 Wednesday 3,112.76
+0.63%



Source: Reuters

By Emily Flitter
Dec. 26, 2018

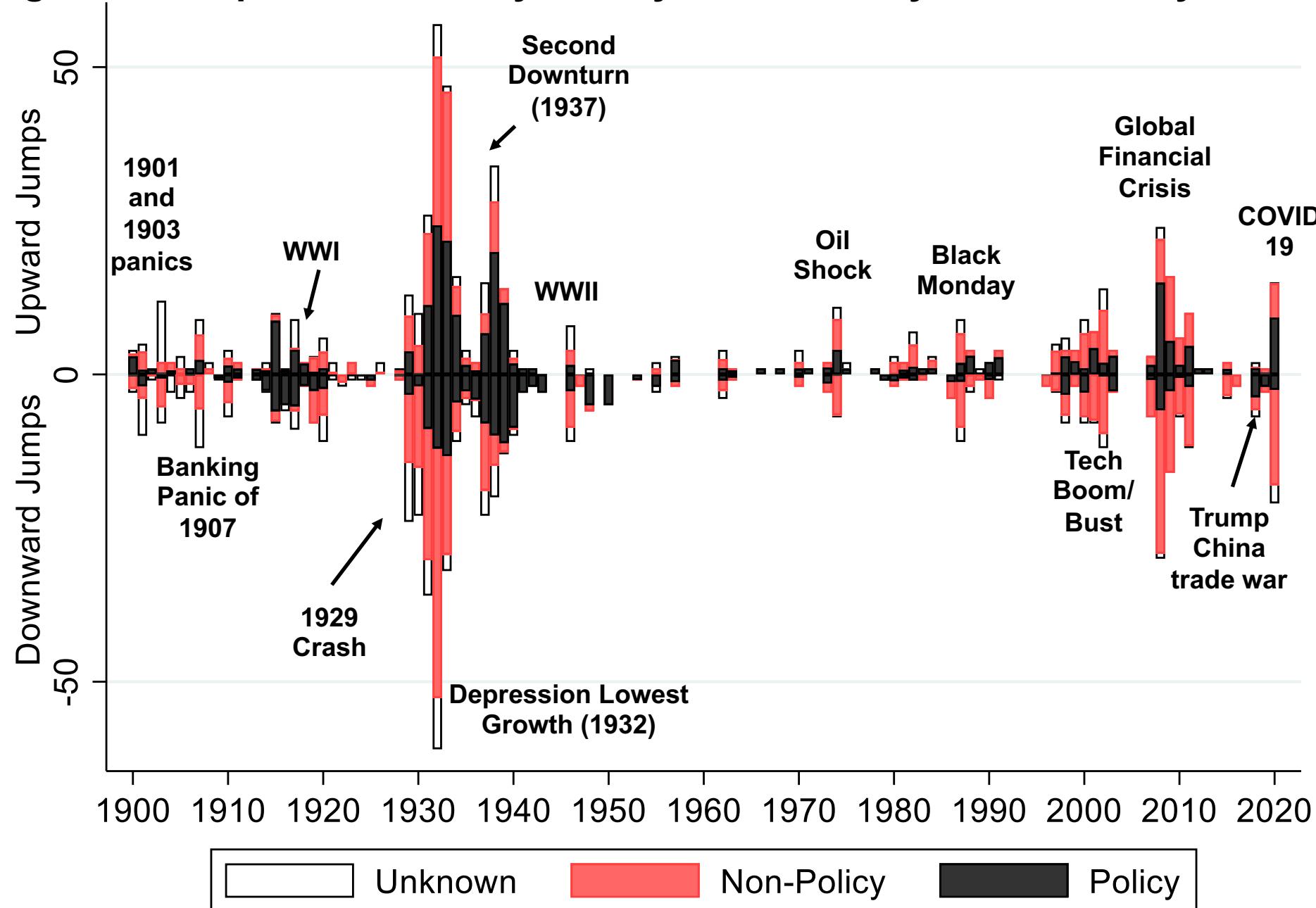
Throughout Wall Street's December meltdown, analysts have been saying that markets were plunging despite plenty of evidence that the United States economy remains strong and corporate profit growth is healthy.

That argument finally found listeners on Wednesday, when early reports of a strong holiday-shopping season helped lift the S&P 500 by nearly 5 percent, its [best day since 2009](#). The Nasdaq added 5.8 percent, and the Dow Jones industrial average rose just under 5 percent. That jump, over 1,086 points, represented the Dow's best single-session gain ever, although a number of days have eclipsed that in percentage terms.

A substantial rise in crude oil prices added to the lighter mood, as did efforts from the White House to ease up on criticism of the Federal Reserve.

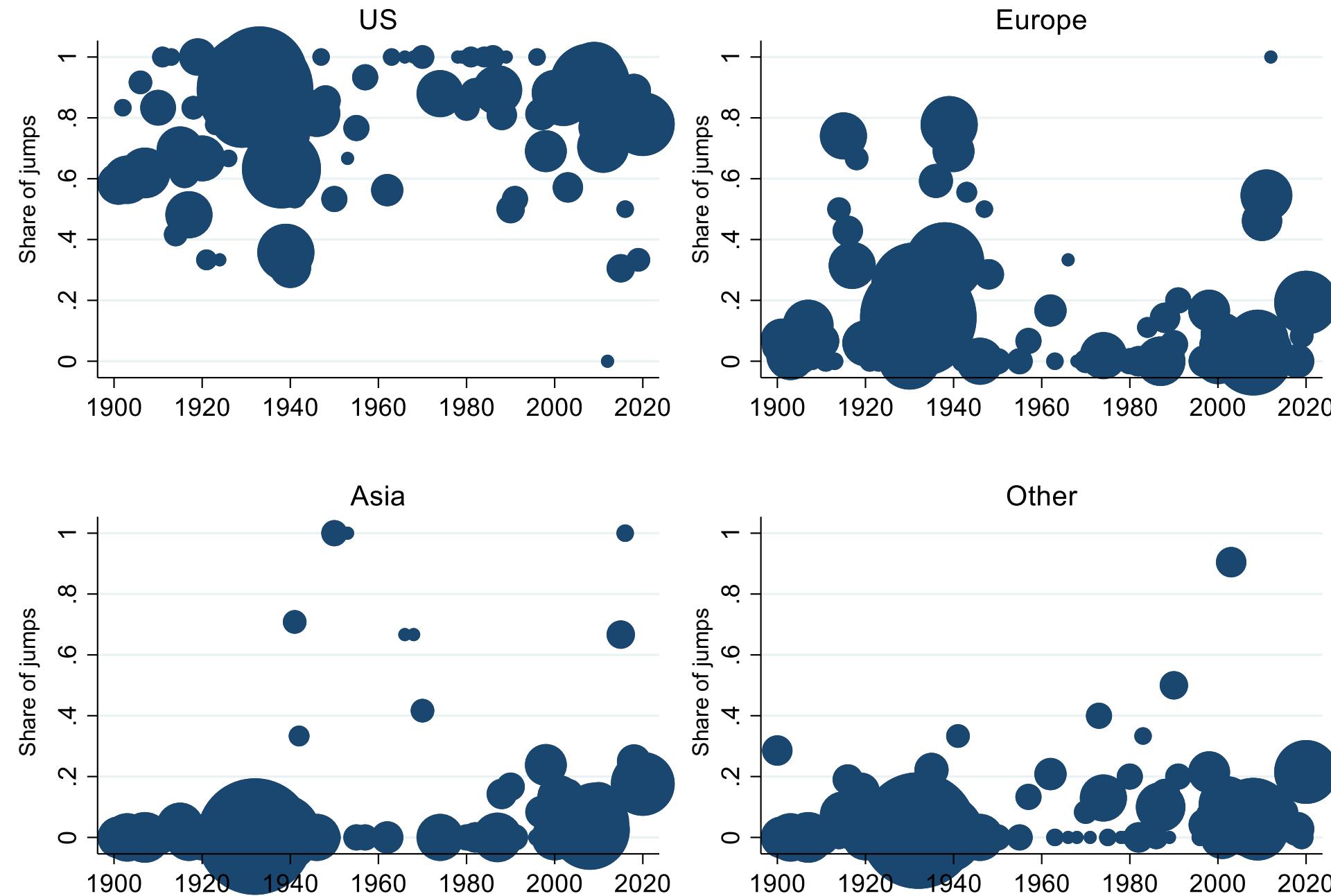
37

Figure 4: Jumps Per Year Vary Greatly but the Policy Share Is Fairly Stable, 1900-2020



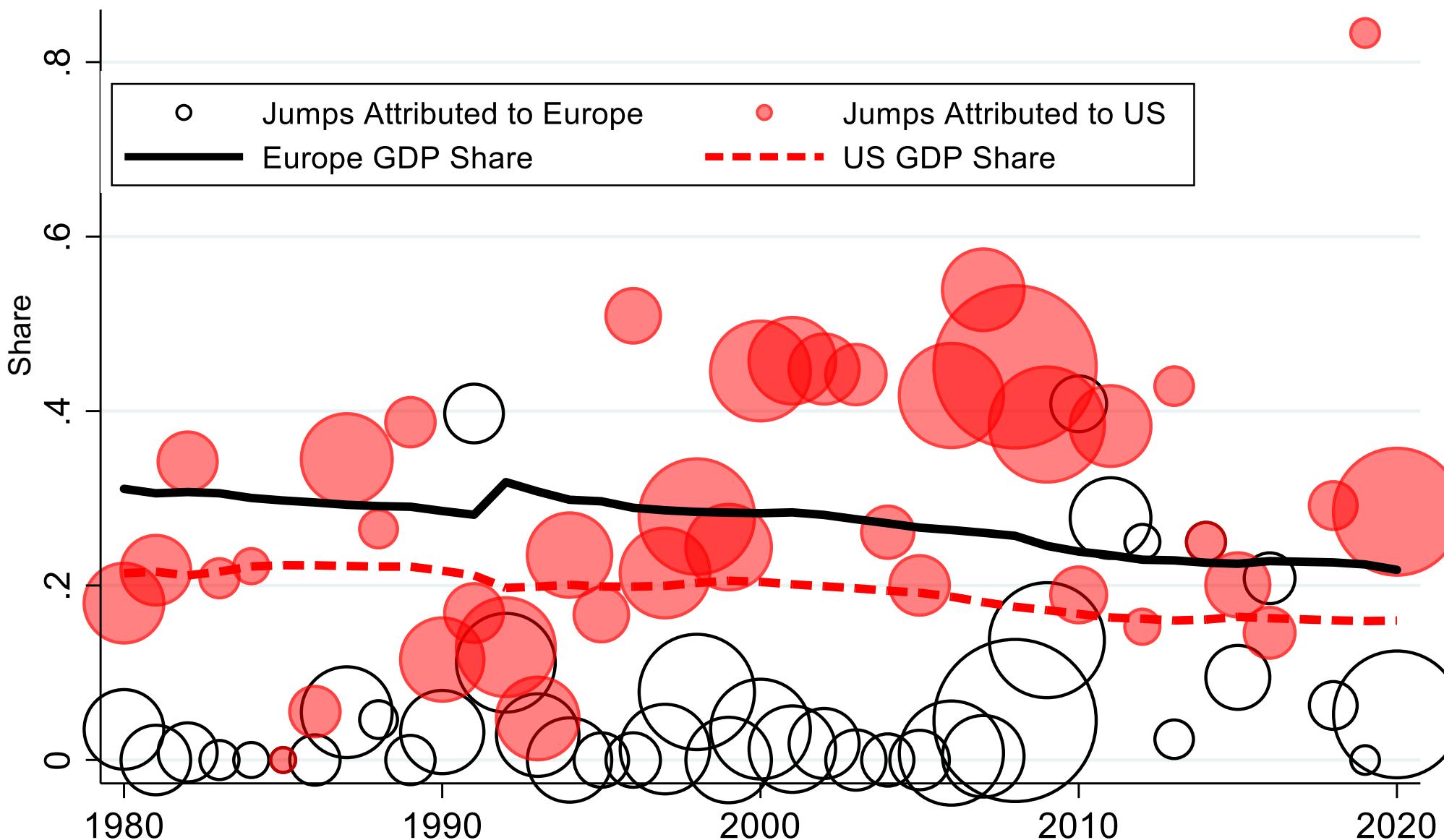
Notes: Each bar is the number of positive or negative jumps in that year. Black and red shadings indicate jumps triggered by "Policy" and "Non-Policy," respectively. The unshaded parts of each bar indicate jumps attributed to "Unknown or No Explanation Offered" plus five instances of "no article found" before 1926.

Figure 5: Jumps in the US Stock Market Are Mostly Due to US News, 1900-2020



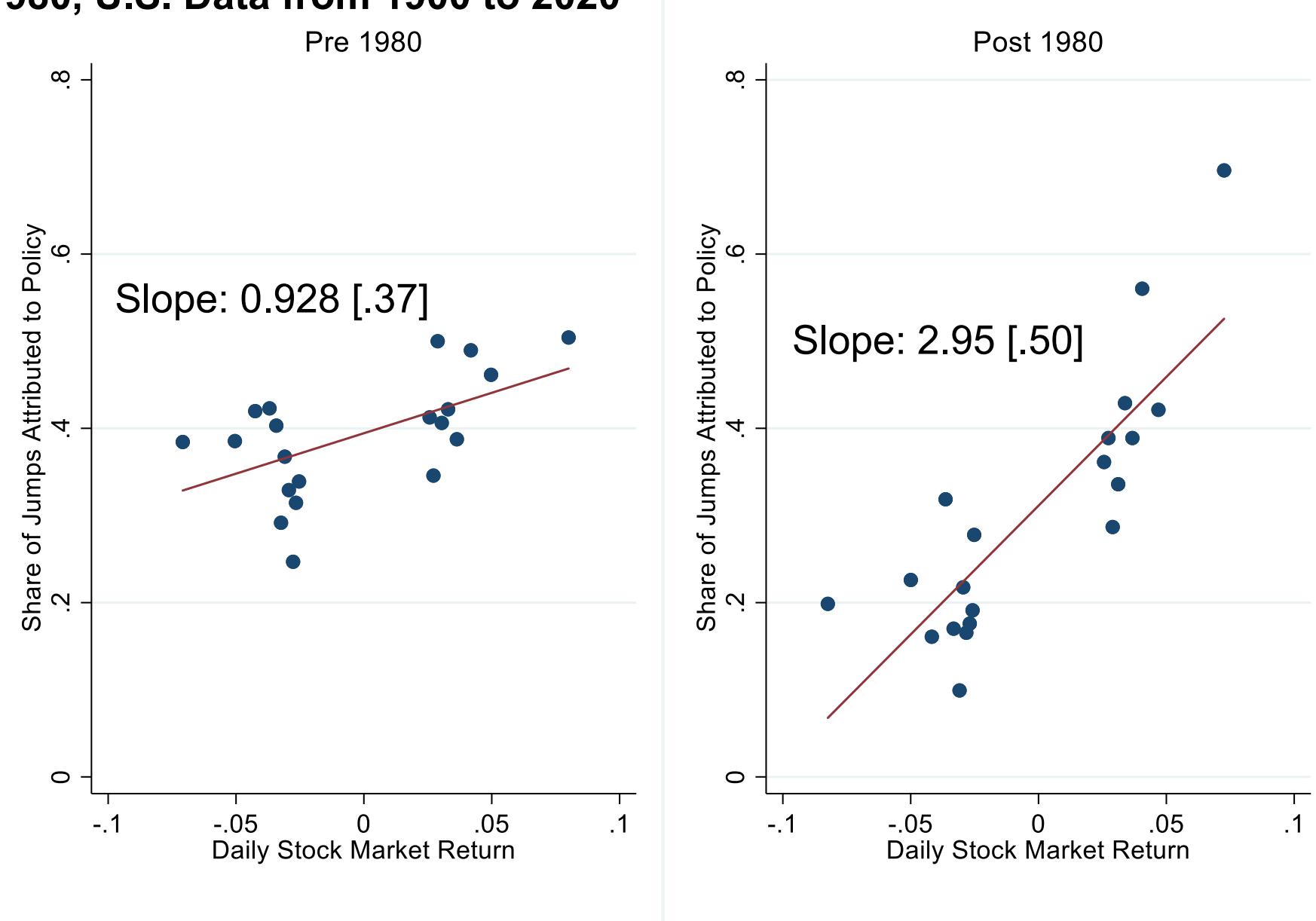
Notes: Dots show the yearly share of U.S. stock market jumps by the geographic origin stated at the top of the panel. Dot size reflects the number of jumps in that year. This chart excludes jumps classified as "Unknown or No Explanation Offered" and "No Article Found," which have no geographic attribution.

Figure 6: News about the United States Triggers a Large Share of National Stock Market Jumps in Other Countries, a Pattern that Does Not Hold for Europe



Notes: This figure shows the yearly share of daily stock market jumps in Australia, Canada, China (HK), China (Shanghai), Japan, New Zealand, Saudi Arabia, Singapore, South Africa, and South Korea that leading own-country newspapers attribute to news about the United States and to news about Europe (or individual European countries). The sample runs from 1980 to 2020 but does not cover all countries in all years. Dot size is proportional to the average number of jumps per country in that year. GDP shares are 5-year moving averages of global GDP, evaluated at PPP, in IMF data. Table A1 reports the sample period by country.

Figure 7: Policy News Triggers A Larger Share of Positive than Negative Jumps, Especially After 1980, U.S. Data from 1900 to 2020



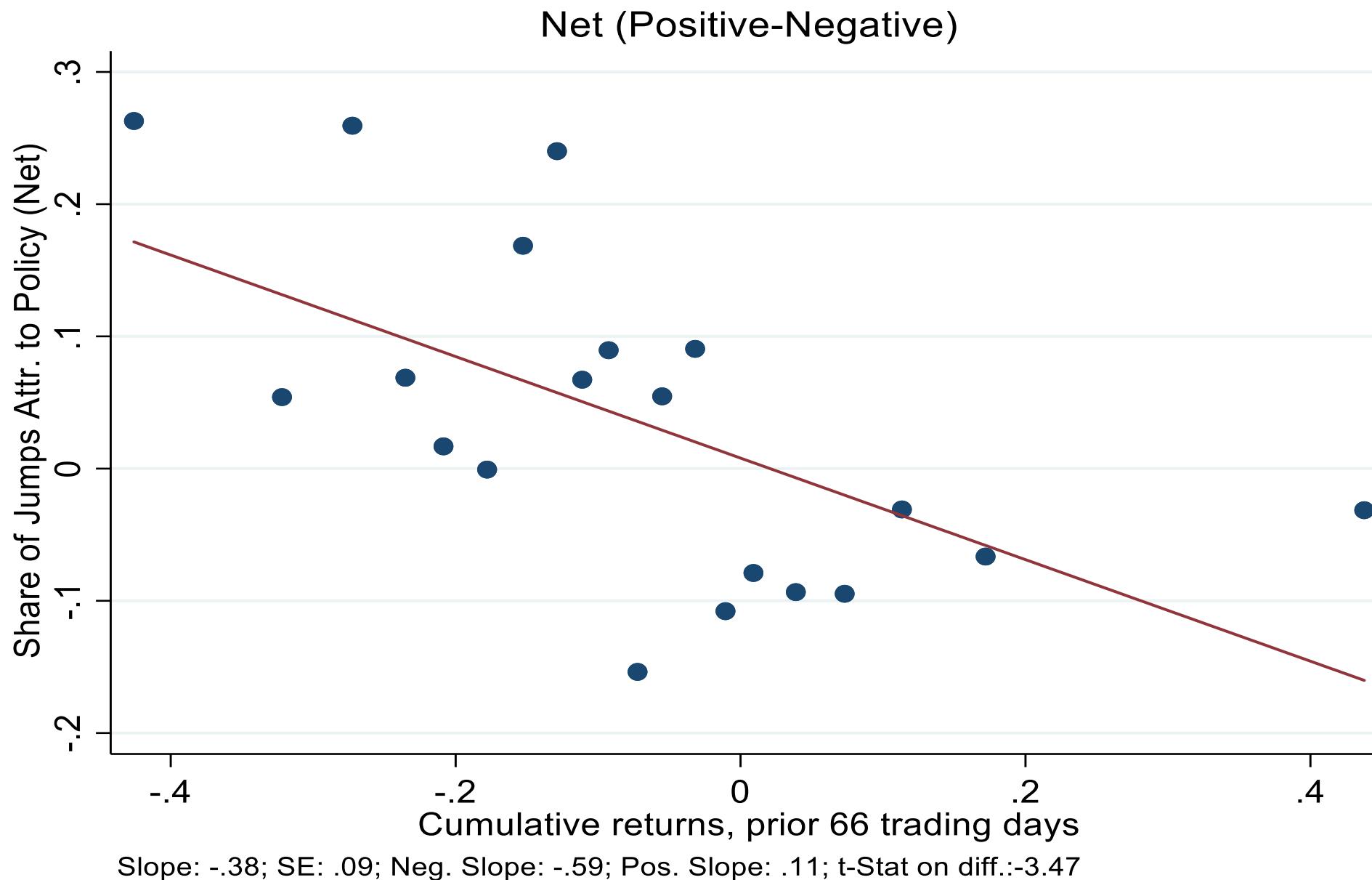
Difference in slopes: 2.03, t-Stat: 3.13.

Notes: Each plot is a binscatter ($n=20$) of jump-level policy scores against jump-day stock returns, where the policy score is the fraction of the jump's codings attributed to policy-related news. For each sub-period, we regress jump-level policy scores on jump-day returns and report the t-statistic on the return variable. For jump days only, we also run the following regression,

$$\begin{aligned} \text{policy}_t \\ = a + b \text{return}_t + c 1_{\text{post}80} \\ + d \text{return}_t \times 1_{\text{post}80} + e_t \end{aligned}$$

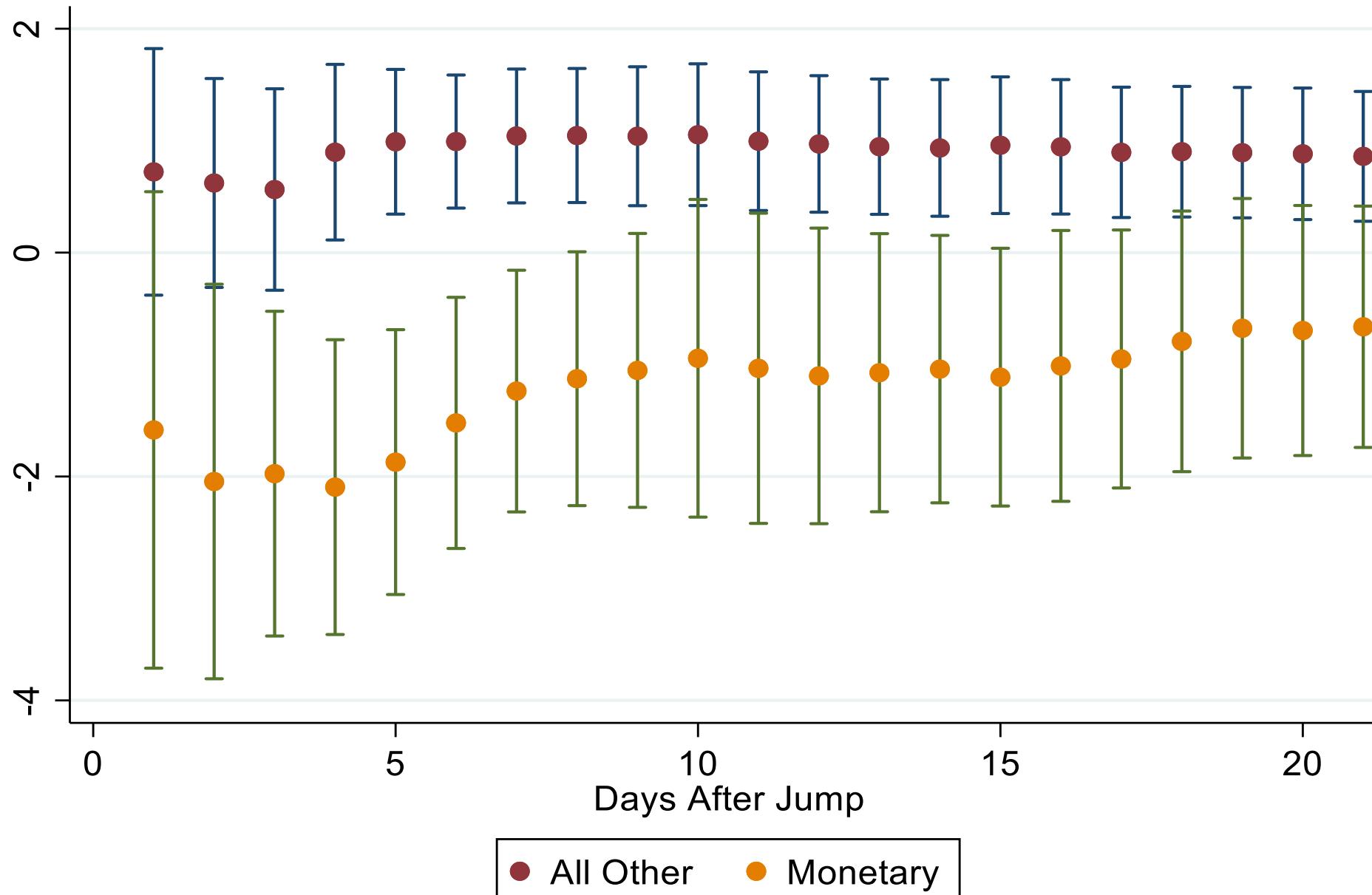
and report the coefficient on the interaction term d , and its t-statistic at the bottom of the figure.

Figure 8: Policy News Is More Likely to Be the Trigger for Positive Jumps When the Market Has Been Falling, U.S. Data from 1900 to 2020



Notes: This chart shows a bin scatter of jump-level net policy scores (on the Y-axis) against the cumulative log returns over the prior 66 trading days. The net policy score equals the share of codings attributed to policy developments for upward jumps, and (-1) times the share of codings attributed to policy for downward jumps.

Figure 9: Volatility Is Lower After Monetary Policy Jumps than After Other Jumps, 1900-2020



Notes: We regress volatility over the next n trading days after a jump day on the fraction of codings that attribute the jump to “Monetary Policy” and the fraction that attribute the jump to “All Other” reasons. We run separate regressions for $n=1, 2, \dots, 22$, and in each case control for the jump-day return, split into positive and negative pieces, and volatility over the day, week and month preceding the jump day (HAR controls). The chart plots coefficients on the two indicator variables and 95% confidence intervals computed using Newey-West standard errors with number of lags set to 1.5 times n .

Figure 9: Volatility Is Lower after Monetary Policy Jumps than after Other Jumps, 1900-2020

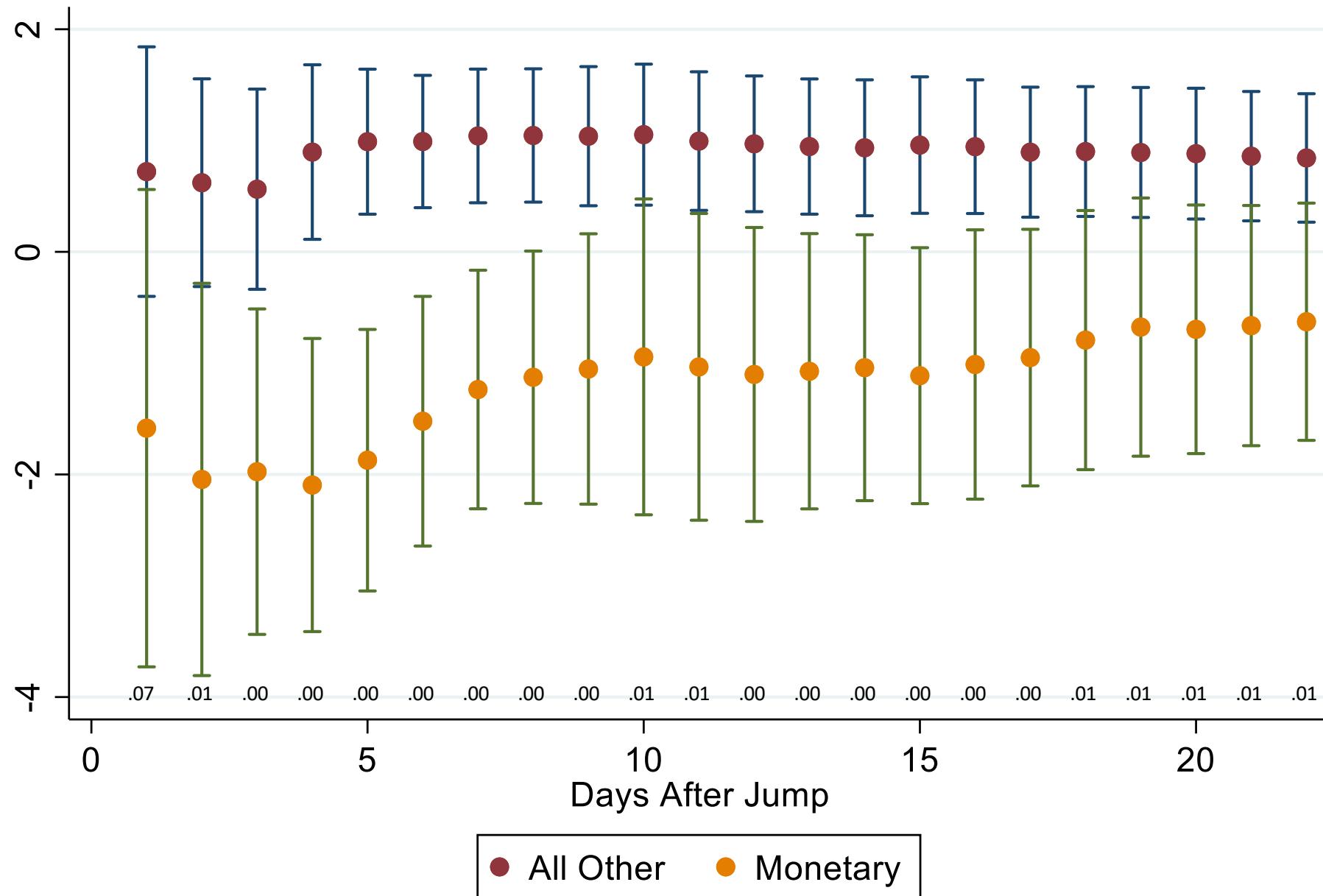
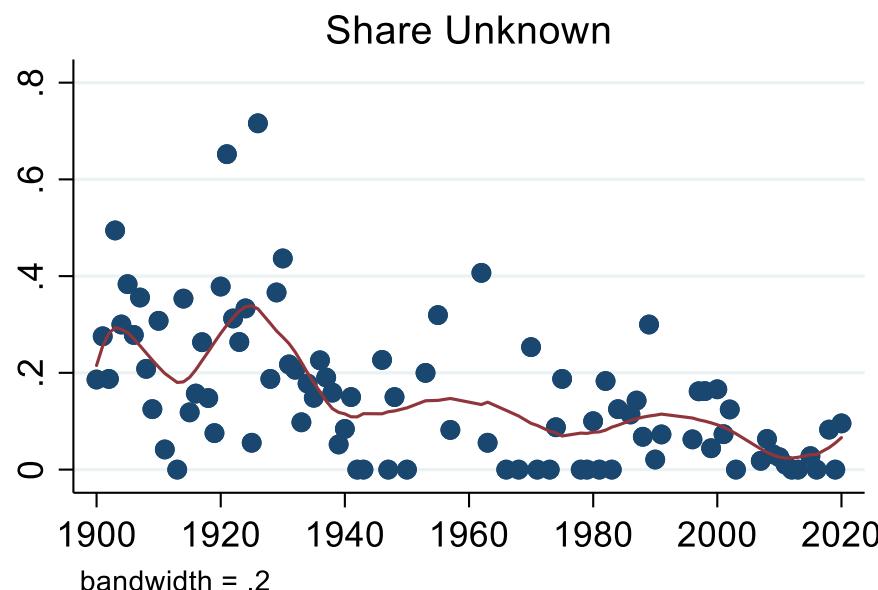
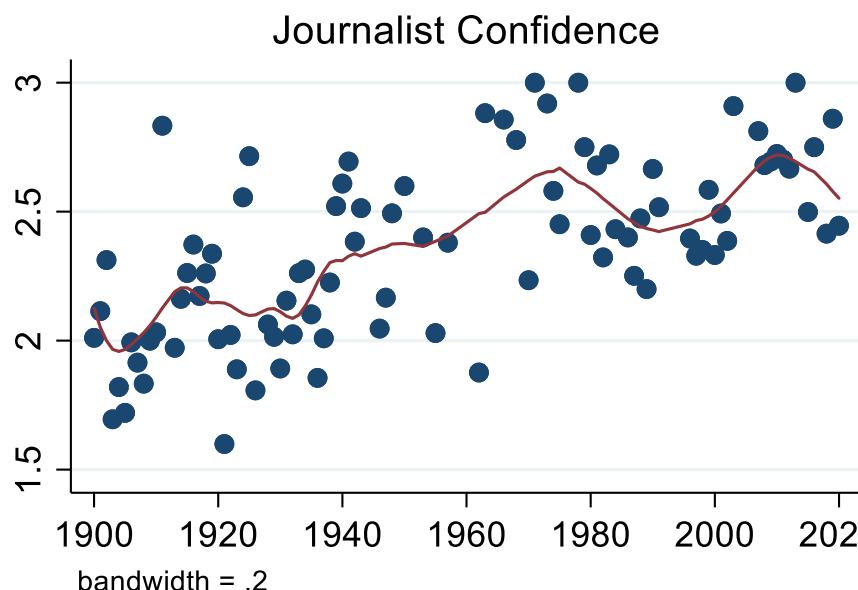
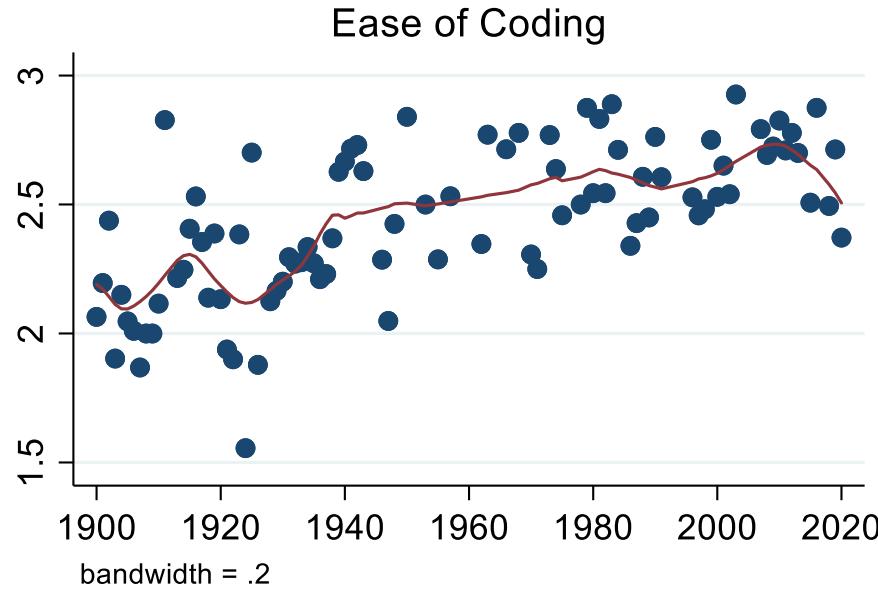
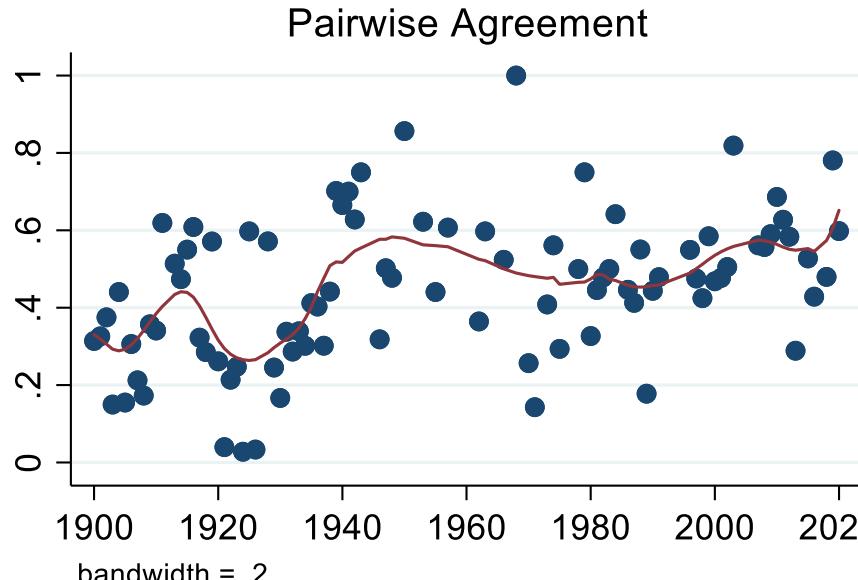
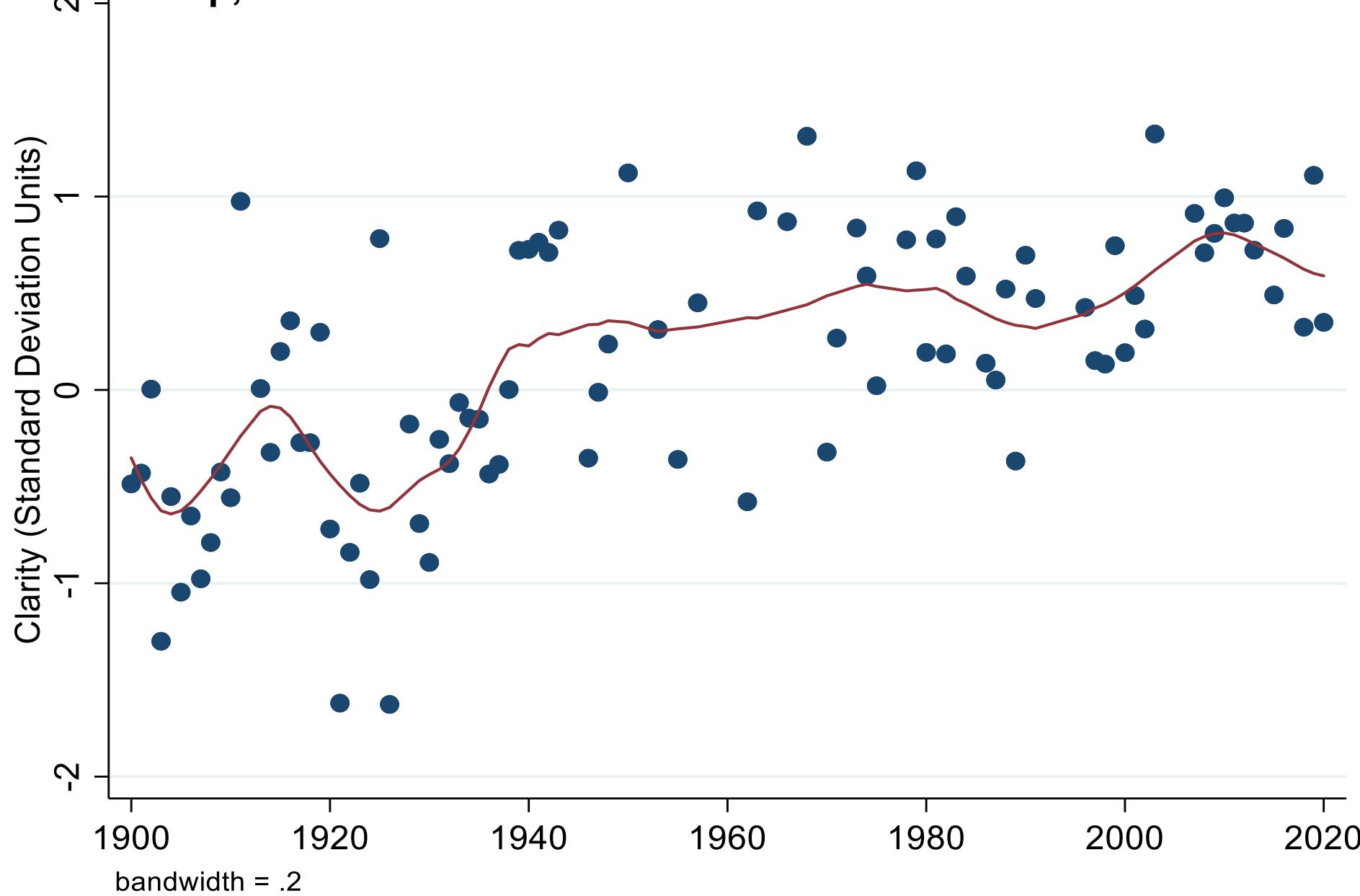


Figure 10: All Four Clarity Index Components Have Trended Towards Greater Clarity, U.S. Data, 1900-2020



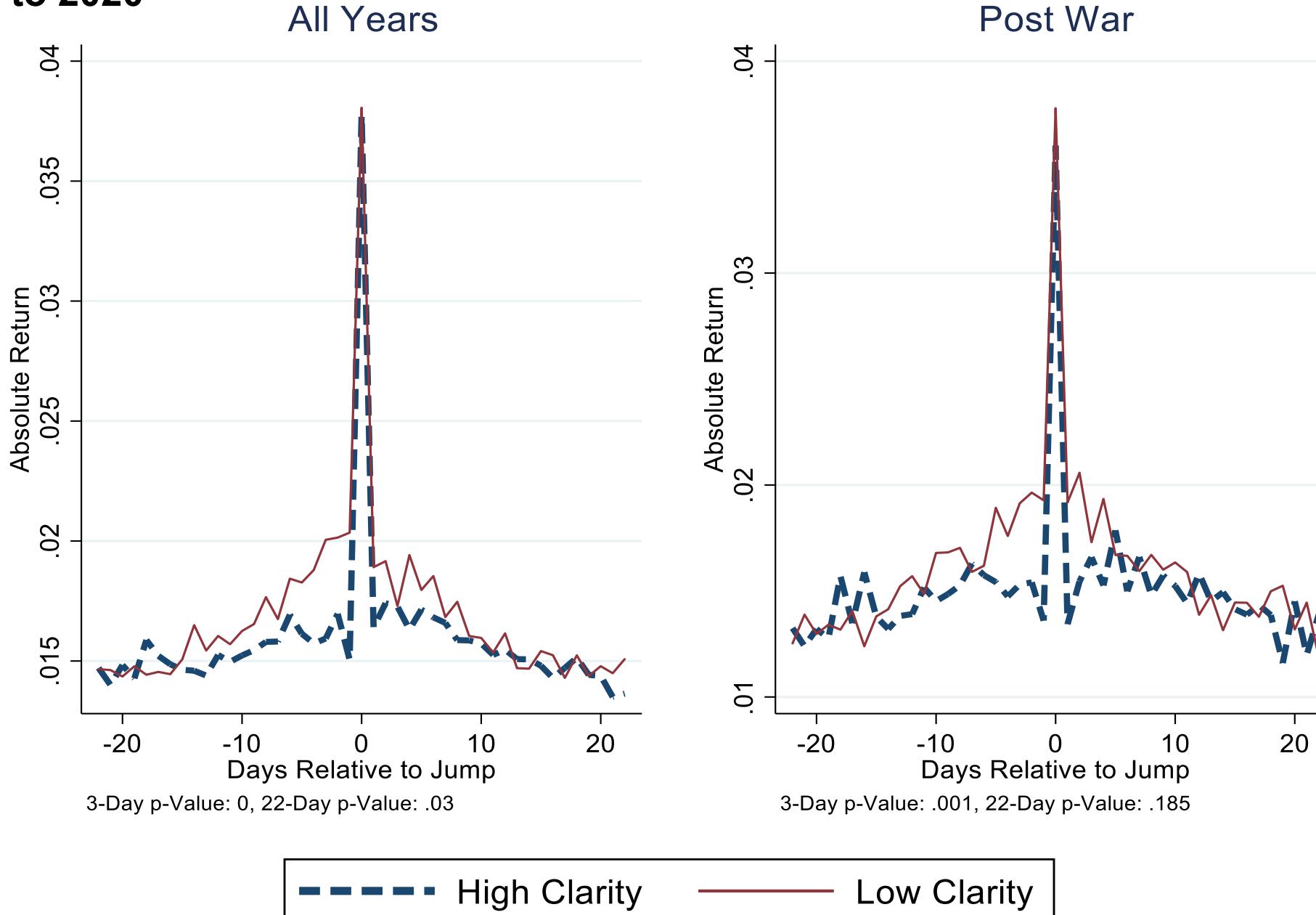
Notes: Each red line shows a LOWESS-smoothed fit to the data, with a bandwidth set to 20 percent of the data. Pairwise Agreement is the average pairwise agreement rate in the codings for a given jump. There are up to 45 pairs arising from 5 newspapers per day and two coders per paper. Ease of Coding is rated on a 1-3 scale, with 3 being the easiest. Journalist Confidence is rated on a 1-3 scale, with 3 being the most confident. Share Unknown is the percentage of codings for a given jump coded as “Unknown or No Explanation Offered.”

Figure 11: Perceived Clarity about the Reason for Daily Stock Market Jumps Has Trended Up, U.S. Data from 1900 to 2020



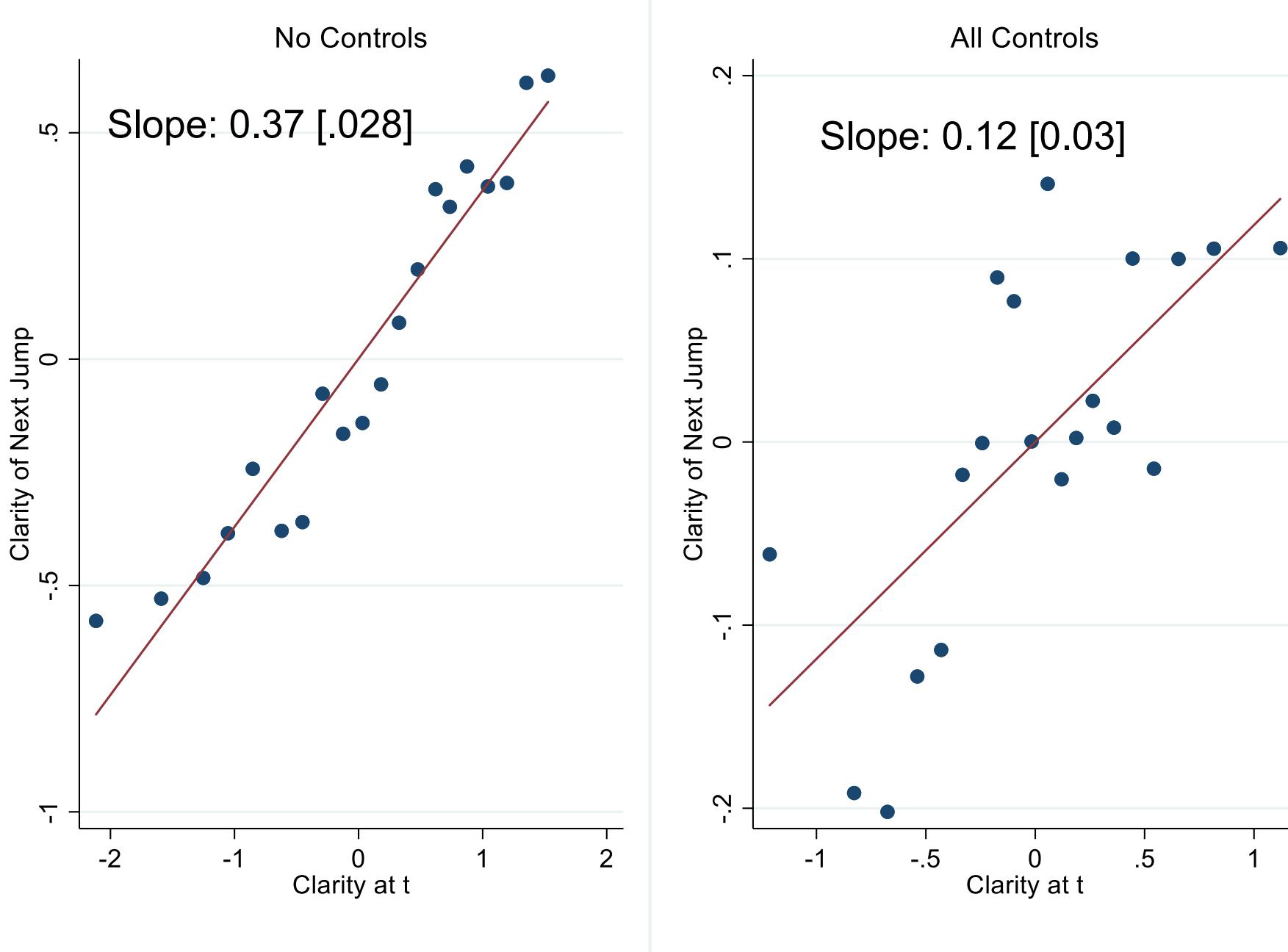
Notes: Clarity is the sum of Ease of Coding, Journalist Confidence, Pairwise Agreement Rate, and the share of codings attributed to “Unknown or No Explanation Offered” (multiplied by -1) after each component is scaled to zero mean and unit standard deviation one. Clarity is also scaled to have zero mean and unit standard deviation. The red line shows a LOWESS-smoothed fit to the clarity data, with a bandwidth set to 20 percent of the data. See the notes to Figure 10 for a description of each Clarity component.

Figure 12: Volatility is Lower Around High-Clarity Jumps, U.S. Data from 1900 to 2020



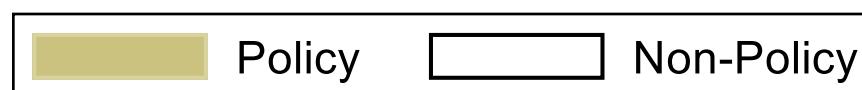
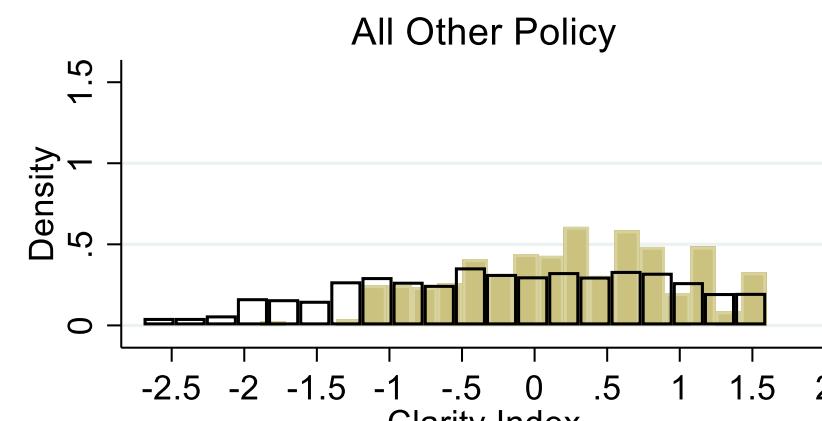
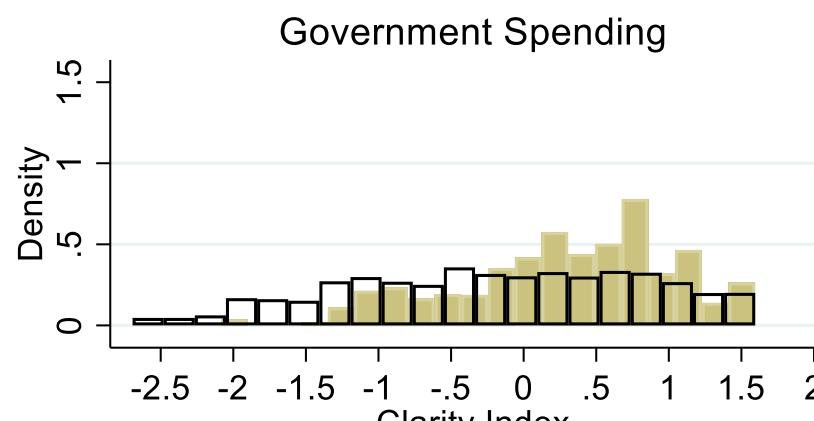
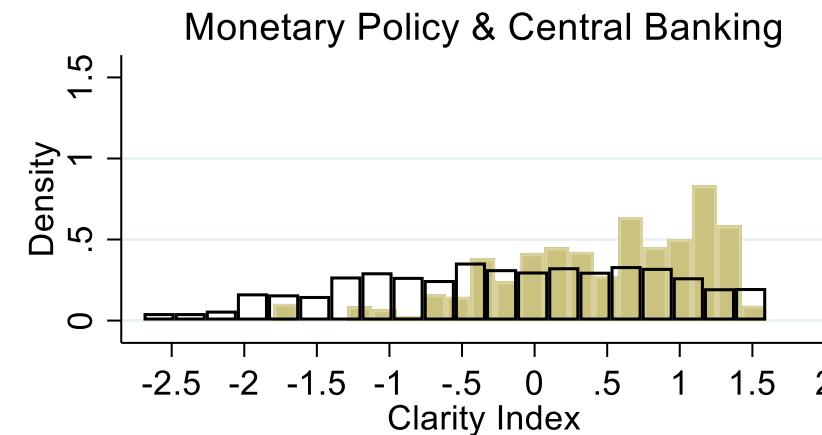
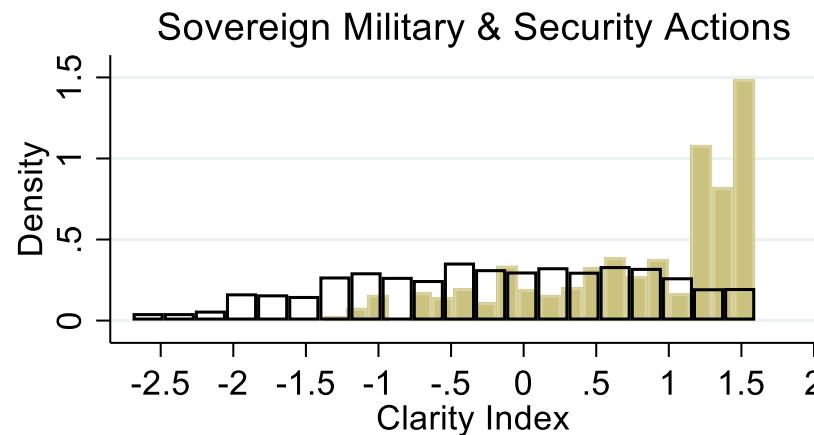
Notes: High (low) clarity is defined as clarity above (below) the sample median for either All Years (1900-2020) or Post War (1946-2020). Each panel shows the average absolute return in a +/- 22-day window around jump days. The p-values are for t-tests of whether the mean absolute return in a +/- n -day window around the jump day differs between high-clarity and low-clarity jumps.

Figure 13: Movements in Clarity Tend to Persist, U.S. Data from 1980 to 2020



Notes: The left panel shows a binscatter ($n=20$) of Clarity of the next jump against Clarity of the current jump. The right panel shows a binscatter using Clarity values that are residualized on the current-day returns, split into positive and negative pieces, and HAR controls (volatility over the day, week and month before the current jump), the 17-way classification of the current jump, the 17-way classification of the next jump, a dummy variable for pre vs. post-World War II, a linear time trend, and the interaction between the time trend and the postwar dummy variable.

Figure 14: Policy Jumps have Higher Clarity Than Non-Policy Jumps on Average, 1900-2020



Notes: The difference in means is the difference in average clarity between each policy category, and all non-policy categories. The t-Stat is from a test of equal means. The number of jumps is the number of codings in each of the policy categories. Non-policy does not include unknown jumps. US data, 1900-2020. Average clarity is higher in every policy subcategory than the average for all non policy subcategories.

Table 1: Distribution of (Primary) Jump Reasons by Era and Category

	US Equities		UK Equities	ROTW Equities	US Bonds	Notes: Thresholds for a day's stock market movements to be considered a 'jump' are listed in Table A1. Jumps are generally calculated for movements of the broadest composite index for a given country. Rest of the World (ROTW) countries include: Australia, Canada, China (HK), China (Shanghai), Germany, Greece, Ireland, Japan, New Zealand, Saudi Arabia, Singapore, South Africa, South Korea and Spain. ROTW panel is not balanced between 1980 and 2020 (see Appendix Table 1). Data for US/UK stock jumps ends 2020. US bond jumps are defined as daily changes in the 10-year treasury yield of more than 15 basis points. This table reports the frequency distributions of the primary reason for the jump, according to newspaper articles. In practice, we often code a secondary (or even tertiary) reason as well, based on the article's explanation for the jump.
	1900-2020	1946-2020	1930-2020	1980-2020	1970-2020	
Time Period:						
Macroeconomic News & Outlook	23.38	32.39	29.12	25.88	53.92	
Corporate Earnings & Outlook	10.99	12.33	10.94	9.23	0.93	
Sovereign Military & Security Actions	9.48	5.25	3.73	2.64	0.90	
Monetary Policy & Central Banking	7.38	10.17	9.53	8.51	18.23	
Government Spending	6.72	6.92	7.14	5.59	4.11	
Commodities	5.57	2.47	2.26	3.45	1.68	
Regulation	4.26	2.23	4.59	2.67	0.27	
Other Non-Policy	4.29	5.94	3.82	10.31	2.80	
Elections & Political Transitions	2.42	3.43	2.78	1.93	0.88	
Other Policy	2.45	1.61	2.15	4.35	0.84	
Taxes	1.71	1.25	0.86	0.49	1.21	
Exchange Rate Policy & Capital Controls	1.07	0.75	0.86	0.66	0.18	
International Trade Policy	0.91	1.31	0.33	0.50	0.08	
Foreign Stock Markets	1.01	0.89	5.98	8.84	0.19	
Terrorist Attacks & Non-State Violence	0.47	0.96	0.89	1.04	0.11	
Unknown & No Explanation	17.44	12.11	12.14	5.86	8.64	
No Article Found	0.43	0.00	2.87	8.05	5.02	
Total	1,152	440	656	4,430	455	

Table 2: Categorical Coding Agreement Rates, U.S. Jumps

Time Period	1900-1945		1946-2020	
	Policy vs. Non-Policy	Granular Categories	Policy vs. Non-Policy	Granular Categories
All Coders & All Papers	75%	43%	80%	54%
All Coders Within Paper	89%	70%	90%	73%
Within WSJ	92%	77%	92%	77%
With Random Assignment	53%	13%	56%	17%
Bootstrapped Std. Err.	10%	7%	12%	10%

Notes: Granular categories include all 17 detailed jump-day categories, including no article found. Policy jumps include Monetary Policy, Government Spending, Sovereign Military, Other Policy, Regulation, Trade Policy, Exchange Rate Policy, Elections, and Taxes. Newspapers include the Wall Street Journal, the NY Times, the Chicago Tribune, the Washington Post, and the LA Times. For the random assignment by period, we use the unconditional distribution of jumps for that sub-period. There are 712 from 1900-1945, and 437 jumps from 1946-2020. Difference between random assignment agreement and all human agreement measures is significant at the 1% level, where standard errors were bootstrapped using 10,000 simulations from the unconditional categorical distribution using all newspapers.

Table 3: Categorical Validation

	Monetary 1994-2020	Macro 1953-2020	Elections 1900-2020	Monetary 1994-2020	Macro 1953-2020	Elections 1900-2020
FOMC meeting at t or t-1	2.144*** (0.261)			2.090*** (0.266)		
Macro Announcement at t	-0.201 (0.156)	0.824*** (0.166)		0.061 (0.226)	0.863*** (0.179)	
Election at t or t-1	-0.485 (1.077)	0.648 (0.965)	4.783*** (0.225)	-0.643 (1.080)	0.662 (0.966)	4.785*** (0.225)
Observations	6,799	17,119	32,787	6,799	17,119	32,787
R-Squared	0.01	0.001	0.014	0.011	0.002	0.014
# Codings in Category	30	137	28	30	137	28
Day of Week FE	No	No	No	Yes	Yes	Yes

Notes: Each column (1) to (3) reports a regression of jump coding values (times 100) for the indicated category on a set of known information-release dates. Because FOMC meetings span two days, we consider jumps that occur on either day of the meeting, or the day after the meeting. We have 259 known FOMC meetings between 1994 and 2020. For elections, because the results are not usually known by the end of the trading date, we consider the day after elections as well. We have 61 known federal elections between 1900 and 2020. For macro news announcements, because they usually occur before the markets open, we only count the day of the announcement. For Macro Announcements, we include the release of CPI, jobless claims and the Employment Situation Report (unbalanced panel between 1953-2020), for a total of 2,732 known dates. Years vary by column. All specifications include day of the week fixed effects. US data, 1900-2020. *** p<0.01, ** p<0.05, * p<0.1. US data, date range varies by column.

Table 4: Policy Jumps are more often positive (negative) after recent stock market falls (rises)

	Positive Policy			Negative Policy			Pos. Policy	Non-Pos. Policy
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cumulative								
Return Past 66 Trading Days	-0.0856	-0.0397	-0.441**	0.214***	0.177***	0.504***	0.133**	-0.0522
	(0.057)	(0.064)	(0.171)	(0.055)	(0.062)	(0.128)	(0.062)	(0.067)
Observations	1,043	603	440	1,043	603	440	1,043	1,043
R-squared	0.018	0.024	0.027	0.032	0.031	0.047	0.017	0.01
Sample:	All	Pre-War	Post-War	All	Pre-War	Post-War	All	All

Notes: For columns 1-3, left hand side variable is equal $policy_t$ if $return_t > 0$, and is zero otherwise. For columns 4-6, left hand side variable is equal $policy_t$ if $return_t < 0$, and is zero otherwise. For column 7, left hand side variable is equal to $nonpolicy_t$ if $return_t > 0$, and zero otherwise. For column 8, left hand side variable is equal to $nonpolicy_t$ if $return_t < 0$, and zero otherwise. In both columns 7 and 8, non-policy excludes jumps classified as unknown. Sample is all US jump days, 1900-2020. Cumulative return past 66 trading days is the cumulative log return on the broad US market index over the past 66 trading days. All specifications include HAR controls: volatility over the past day, week, month and quarter. Pre-war is data from 1900-1937, and postwar is data from 1946-2020.

Table 5: Clarity is Higher When Intraday Movement is More Concentrated

	Concentration x 100					
	(1)	(2)	(3)	(4)	(5)	(6)
Clarity	2.179*** (0.584)	1.97*** (0.610)				
Avg. Ease of Coding			0.99* (0.576)			
Avg. Confidence				1.88*** (0.578)		
Share Unknown					-1.82** (0.799)	
Pairwise Agreement						1.78*** (0.449)
Observations	322	322	322	322	322	322
R-squared	0.05	0.088	0.059	0.086	0.074	0.096
Return Controls	NO	YES	YES	YES	YES	YES
HAR Controls	NO	YES	YES	YES	YES	YES

Notes: Clarity has mean zero and standard deviation one. The left hand side variable is share of total distance traveled in the 15 minute window with the largest absolute return, multiplied by 100. Sample spans US data for which high frequency data is available from TickWrite for the S&P 500 Spot Market, 1985-2020. *** p<0.01, ** p<0.05, * p<0.1

Table 6: Higher Clarity Jumps are Followed by Less Returns Volatility and Cross-Firm Dispersion

	Volatility			X-Sectional Std. Dev.		
	(1)	(2)	(3)	(4)	(5)	(6)
Clarity	-4.975*** (1.57)	-4.846*** (1.38)	-2.232* (1.25)	-0.296*** (0.05)	-0.270*** (0.04)	-0.0997*** (0.03)
Observations	1,146	1,146	1,146	921	921	919
R-squared	0.009	0.156	0.248	0.043	0.187	0.564
Controls	None	Returns	+HAR	None	Returns	+HAR
Sample	All	All	All	All	All	All

Notes: For columns 1-3, left-hand-side is the sum of squared percentage returns over the 5 days following a jump day. US data, 1900-2020. For columns 4-6, left-hand-side is the average value-weighted cross-sectional standard deviation over the 5-days following the jump, multiplied by 100. This cross-sectional standard deviation is computed using all ordinary common shares traded on major exchanges in CRSP. US data, 1926-2020Q3. Column 6 has 2 fewer observations because we are including lagged cross sectional standard deviation on the right-hand-side. For columns 2 and 4 controls are the jump day return, split into positive and negative components. For column 3, controls also include: 1-day, 5-day and 22-day lagged volatility (HAR controls). For column 6, controls also include: 1-day, 5-day and 22-day lagged cross-sectional standard deviation (i.e. HAR for cross-sectional standard deviation). Clarity has mean zero and standard deviation one. Robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

Table A1: Countries, Newspapers and Jump Thresholds

Country	Years	Sources	Jump Threshold
United States	1900-2020	Wall Street Journal, etc.	2.50%
United Kingdom	1930-2020	Financial Times (UK Edition), Times of London, Telegraph, etc.	2.50%
Australia	1986-2020	Australian Financial Times, Sydney Morning Herald	2.50%
Canada	1980-2020	The Globe and Mail	2.00%
China (Hong Kong)	1989-2011	South China Morning Post	3.80%
China (Shanghai)	1993-2013	Shanghai Securities News, China Securities Journal	4.00%
Germany	1987-2020	Handelsblat, FAZ	2.50%
Greece	2001-2015	Kathimerini, To Vima	4.00%
Ireland	1987-2012	The Irish Times	2.50%
Japan	1981-2013	Yomiuri, Asahi	3.00%
New Zealand	1996-2011	New Zealand Herald	2.50%
Saudi Arabia	1994-2013	Al Riyadh	2.50%
Singapore	1980-2020	Business Times, Straits Times	2.50%
South Africa	1986-2020	Business Day	2.50%
South Korea	1981-2011	Chosun Ilbo	3.50%
Spain	1987-2020	ABC	2.50%

Notes: The jump threshold is the minimum absolute return required for a day to be considered a jump in each country. We allow for differences across countries to account for differences in unconditional volatility.

Table A2: Regression Models Fit to Daily Industry-Level Equity Returns from 1960 to 2016

	<i>Banks</i>	<i>Pooled Sample</i>		
	(1) All Days	(2) Jump Days	(3) All Days	(4) Jump Days
γ Coefficient	0.80***	0.74***	0.55***	0.51***
(St. Error)	(0.23)	(0.24)	(0.13)	(0.13)
Observations	13,469	339	109,760	4720
R-Squared	0.67	0.83	0.56	0.81

Notes: See Appendix A for the regression specification and the interpretation of the γ coefficient. We use Fama-French industry-level returns data. A single-industry regression for Guns, yields results similar to the Pooled Sample, but the standard error is large and the coefficient estimate is insignificant. When we set $Tri=-1$ for the Aerospace industry for jumps attributed to Sovereign Military Conflict, the Aerospace regression yields a small, marginally significant coefficient of the wrong sign. That may reflect the ambiguous nature of Aerospace firms' responses to military conflict: (relatively) good news for defense-oriented aerospace firms may, at the same time, be bad for aerospace firms oriented toward civilian customers. If we set $Tri=1$ for Aerospace in jumps attributed to Sovereign Military Conflict, the anomalous Aerospace result disappears, and the Pooled Sample results get stronger.
 *** p<0.01, ** p<0.05, * p<0.1

Table A3: Comparison to the Cutler, Poterba and Summers Characterization of the 50 Largest Daily Moves in the S&P Stock Index from 1946 to 1987

	Primary or secondary category agreement	Primary category agreement	Observations
<i>New York Times</i>			
High clarity	81%	75%	32
Low clarity	43%	38%	18
Total	67%	62%	50
<i>Wall Street Journal</i>			
High clarity	64%	55%	32
Low clarity	48%	37%	18
Total	58%	48%	50

Notes: Cutler, Poterba and Summers (CPS) attribute a “cause” to the 50 largest U.S. stock market jumps from 1946 to 1987 based on coverage in the New York Times. See their Table 4. For each jump, we map their description of the cause to a primary category and, sometimes, a secondary category, according to our classification scheme. We then compare the resulting CPS classifications to our classification as follows: For any given coding of the jump in question, we set “Primary category agreement” to 1 if the CPS primary category matches ours, and 0 otherwise. We set “Primary or secondary category agreement” to 1 if there is any overlap between the CPS primary and secondary categories and our primary and secondary categories, and 0 otherwise. We then average over all codings for the jump in question to obtain an average agreement rate (over codings) per jump. Lastly, we average over jumps to obtain the entries reported in the table. “High” and “Low” clarity jumps have Clarity values greater or less than 0, respectively.

Table A4: Policy-Share by Jump Size and Period, US

Absolute Jump Size	US				Rest of the World	
	1900-1979		1980-2020		1980-2020	
	Positive	Negative	Positive	Negative	Positive	Negative
[Thresh,Thresh+0.5%)	41%	31%	35%	20%	32%	21%
[Thresh+0.5%,Thresh+1%)	41%	34%	38%	12%	42%	22%
[Thresh+1%,Thresh+1.5%)	40%	42%	46%	34%	29%	22%
Thresh+1.5% or Larger	48%	40%	55%	19%	47%	24%
All	45%	33%	43%	20%	41%	23%
p-Value	0.01		0.00		0.00	
Total	802		350		3,664	

Notes: Positive (Negative) columns are share of positive (negative) jumps attributed to policy categories. For rest of the world, we exclude jumps attributed to Unknown or No Article Found when computing the totals. p-Value is from a t-Test that share of policy-share is the same among positive and negative jumps. US data 1900-2020.

Table A5: Policy-Share by Jump Size, Period and Category, US

	1900-1979		1980-2020		Post-1980		p-Value
	Positive	Negative	Positive	Negative	Return Shift		
Policy	161	152	71	36	0.023	0.000	
Sovereign Military & Security Actions	33	65	5	6	0.011	0.401	
Monetary Policy & Central Banking	31	16	29	9	0.014	0.152	
Government Spending	36	12	19	10	-0.009	0.498	
Regulation	21	25	2	1	0.004	0.888	
Taxes	7	9	4	0	0.042	0.007	
All Other Policy	33	26	11	9	0.005	0.693	
Non Policy	136	187	76	132	-0.024	0.000	
Macroeconomic News & Outlook	68	79	43	79	-0.026	0.000	
Corporate Earnings & Outlook	33	44	23	26	0.003	0.786	
Commodities	24	34	2	4	-0.017	0.435	
All Other Non-Policy	11	29	9	22	0.000	0.995	

Notes: Results shown for most common policy and non-policy categories. The final two columns are from the coefficient d in the regression:

$$return_t = a + b \text{ category}_t + c 1_{post80} + d \text{ category}_t \times 1_{post80} + e_t$$
US data, 1900-2020.

Table A6: Breakdown of Policy Jumps by Country

	Non-Policy		Monetary Policy		Government Spending		All Policy	
	Negative	Positive	Negative	Positive	Negative	Positive	Negative	Positive
Australia	84	42	4	9	5	15	15	30
Canada	206	135	24	28	9	15	45	51
China	118	107	14	7	0	0	58	63
Germany	198	131	15	27	7	14	46	65
Greece	54	34	6	11	10	22	29	42
Hong Kong	87	56	8	22	5	4	23	43
Ireland	148	106	8	16	7	7	42	56
Japan	133	94	6	17	7	15	35	61
South Korea	144	128	5	12	0	5	36	64
New Zealand	33	16	0	1	0	1	0	2
Saudi Arabia	151	96	0	1	0	2	16	22
Singapore	145	113	12	16	3	12	28	39
South Africa	165	121	10	20	5	13	32	54
Spain	216	160	24	55	26	39	92	124
UK	240	191	27	44	18	29	100	125
All	2,121	1,530	161	287	102	193	596	839

Notes: Positive (Negative) columns are share of positive (negative) jumps attributed to policy categories. Data from 1980-2020.

Table A7: How the likelihood of positive and negative jumps attributed to monetary and fiscal policy vary with recent market performance

	Positive Monetary			Positive Fiscal		
	(1)	(2)	(3)	(4)	(5)	(6)
Cumulative Return Past 66 Trading Days	-0.0473*	-0.0271	-0.201**	-0.0880***	-0.0647*	-0.170***
	(0.026)	(0.028)	(0.099)	(0.034)	(0.038)	(0.063)
Observations	1,043	603	440	1,043	603	440
R-squared	0.005	0.01	0.01	0.032	0.018	0.057
Sample:	All	Pre-War	Post-War	All	Pre-War	Post-War
	Negative Monetary			Negative Fiscal		
	(7)	(8)	(9)	(10)	(11)	(12)
Cumulative Return Past 66 Trading Days	0.0701***	0.0557**	0.207**	-0.0205	-0.0336	0.0136
	(0.021)	(0.022)	(0.086)	(0.028)	(0.031)	(0.040)
Observations	1,043	603	440	1,043	603	440
R-squared	0.021	0.02	0.038	0.019	0.03	0.015
Sample:	All	Pre-War	Post-War	All	Pre-War	Post-War

Notes: For columns 1-3, the dependent variable equals *monetary policy*_t if *return*_t > 0, and zero otherwise. For columns 4-6, the dependent variable equals *fiscal policy*_t if *return*_t > 0, and zero otherwise. *Fiscal policy* is defined as the sum of government spending and tax policy. The sample is all US jump days, 1900-2020. Cumulative return past 66 trading days is the cumulative log return on the broad US market index over the preceding 66 trading days. All specifications include HAR controls: volatility over the past day, week, month and quarter. Pre-war data run from 1900 to 1937, and postwar data run from 1946 to 2020.

Table A8: Clarity Persistence, US, 1900-2020

	LHS: Clarity of Jump at t (All Years) x 100						
Clarity of Last Jump	0.237*** (0.0320)	0.181*** (0.0438)	0.132*** (0.0311)	0.133*** (0.0308)	0.113*** (0.0314)	0.117*** (0.0312)	0.114*** (0.0390)
Return of Last Jump (Positive)				-404.3*** (121.5)		-306.0** (134.3)	-286.7** (136.3)
Return of Last Jump (Abs, Negative)				-475.9*** (115.1)		-403.6*** (126.8)	-397.8*** (127.0)
Volatility last day					-846.9 (1103.0)	-814.2 (1043.0)	-621.8 (1021.0)
Volatility last week					-668.4 (570.3)	-378.3 (581.6)	-328.1 (582.3)
Volatility last month					-135.2 (203.3)	-124 (202.2)	-78.35 (208.0)
In(Time Since Last Jump)							1.641 (1.4)
In(Time Since Last Jump) x Clarity							0.0592 (1.2)
Observations	1,151	1,151	1,151	1,151	1,149	1,149	1,149
R-squared	0.222	0.256	0.62	0.625	0.625	0.628	0.629
Specification	Add Cat. Of Last Jump	Add Cat of Current Jump	Return Controls	Har Controls	Return/Har controls	In(Time since last jump)	

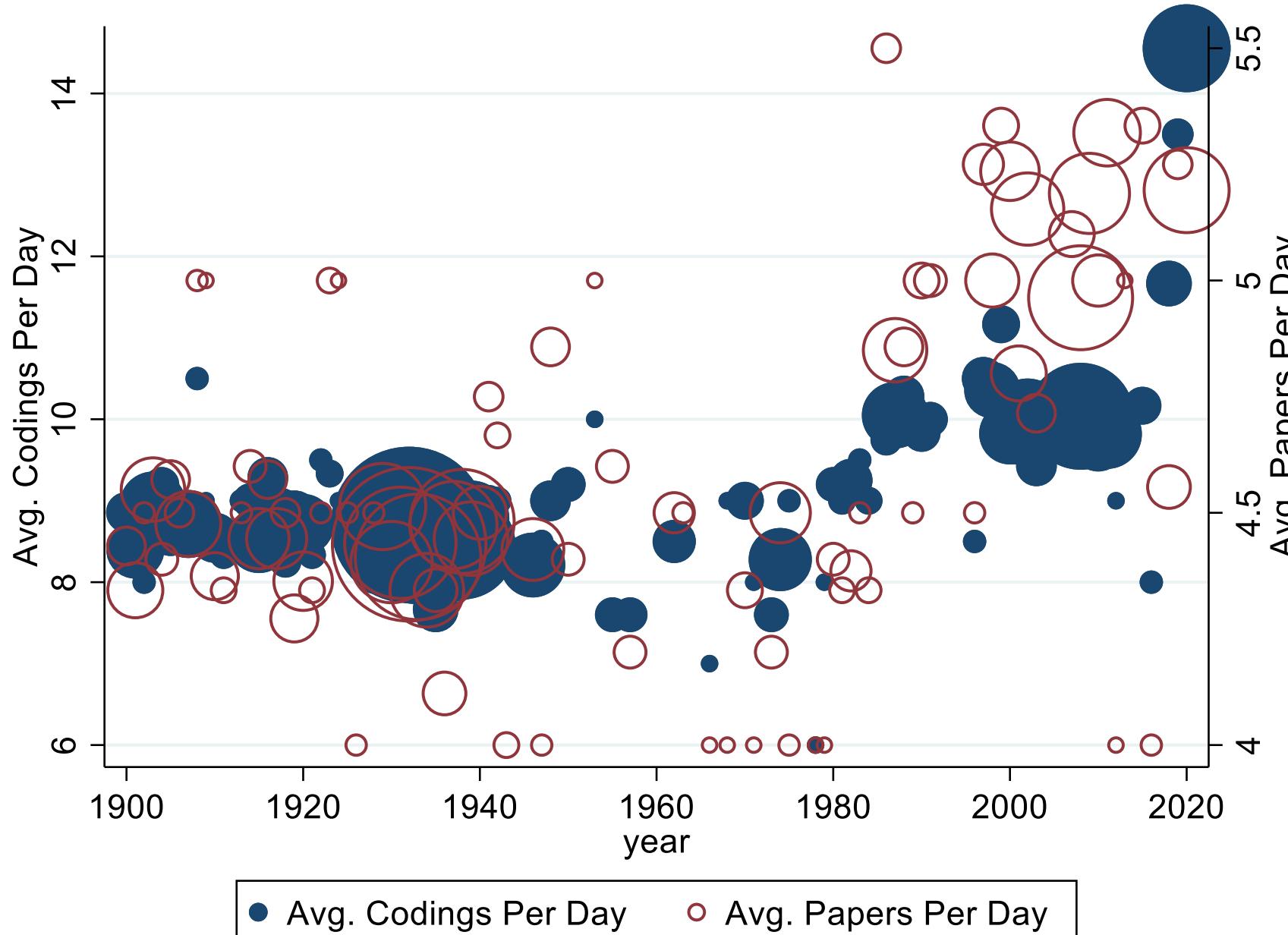
Notes: Regressions, where the left-hand-side is the clarity of a jump on date t (multiplied by 100), and the right hand side was the clarity of the last jump chronologically. The first column has a linear time trend, a postwar dummy variable, and an interaction term between the postwar dummy and the time trend. The second column adds the category of the last jump, while the third column adds the category of the current jump – as long as neither of these categories are unknown. The fourth column adds controls for the return of the last jump, while the fifth column has HAR controls, all relative to the last jump (last 1-day, 1-week and 1-month volatility). The sixth column has both the past return and HAR controls. The 7th column controls for the log of time since the last jump (in days) and interacts this with the clarity of the last jump. US data, 1900-2020. Robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

Table A9: Volatility Following Policy and Non-Policy Jumps, US, 22-day

		Next 22 Days		Next 5 Days			
		(1)	(2)	(3)	(4)	(5)	(6)
	Policy	3.628*** (0.564)	0.26 (0.331)				
	Non-Policy	4.673*** (0.637)	0.991*** (0.361)				
Non-Policy	Commodities		6.954*** (1.210)	1.716* (0.891)	9.185*** (2.219)	2.489* (1.442)	
	Corporate Earnings		3.712*** (0.797)	0.788 (0.538)	3.870*** (1.358)	-0.0762 (0.871)	
	Macro News		4.882*** (0.789)	1.359*** (0.522)	5.965*** (1.140)	1.393 (0.963)	
Policy	Monetary Policy		2.187*** (0.636)	-0.668 (0.556)	1.772*** (0.662)	-1.900*** (0.602)	
	Fiscal Policy		6.591*** (1.559)	1.639 (1.298)	7.299*** (1.919)	0.983 (1.253)	
	Sovereign Military		1.538*** (0.391)	-0.559 (0.375)	3.266*** (0.871)	0.27 (0.786)	
Obs		32,775	32,753	32,775	32,753	32,753	32,753
R-Squared		0.109	0.323	0.117	0.325	0.107	0.312
Return Controls		NO	YES	NO	YES	NO	YES
HAR Controls		NO	YES	NO	YES	NO	YES
F-Test for joint equality of coeffs.		0.0594	0.102	4.94E-05	0.0147	5.04E-05	0.00807

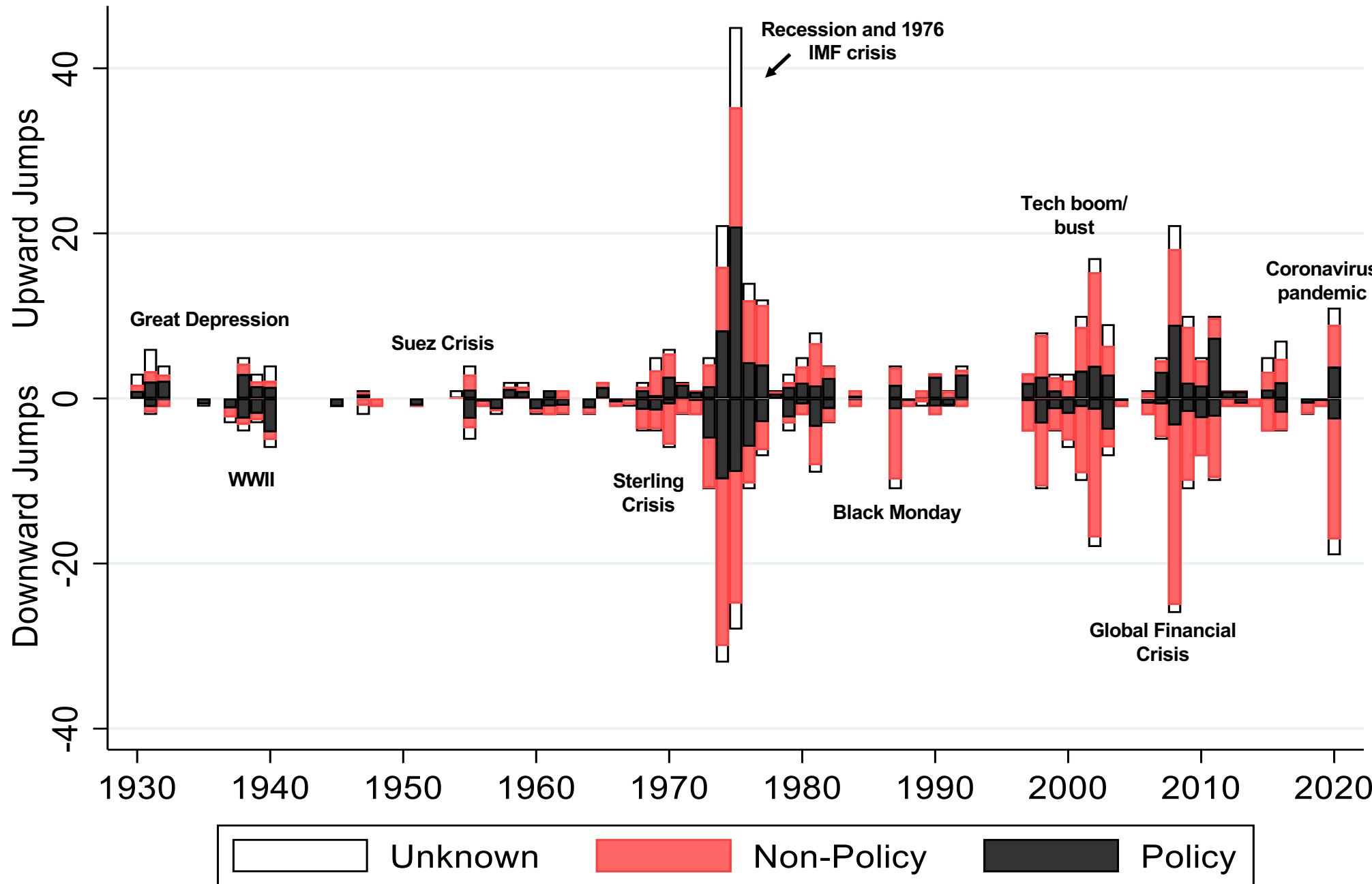
Notes: Columns 1-4 represent regressions, where the left-hand-side is the average percentage squared return in the 22 days following the jump. In columns 5-6, the left-hand-side is the average percentage squared return in the 5 days following the jump. US data, 1900-2020. There are only dummy variables for the jump categories shown, as well as a residual category which includes all the non-enumerated categories. Fiscal policy includes government spending and taxes. Enumerated categories represent the categories with the highest number of jumps by policy/non-policy groups. Columns 1-4: Newey-West standard errors with 33 lags. Columns 1-4: Newey-West standard errors with 8 lags.
 *** p<0.01, ** p<0.05, * p<0.1

Figure A10: Average Number of Codings and Newspapers per Day by Year for U.S. Stock Market Jumps, 1900 to 2020



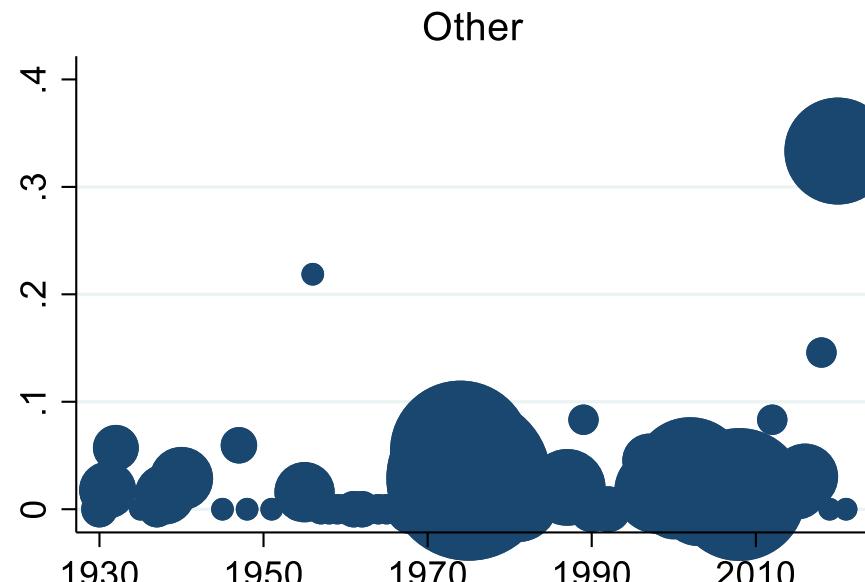
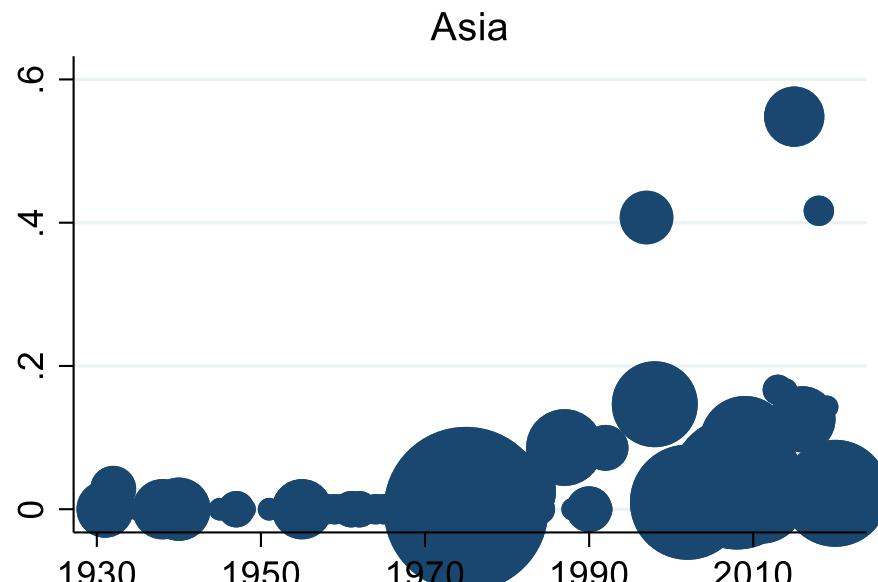
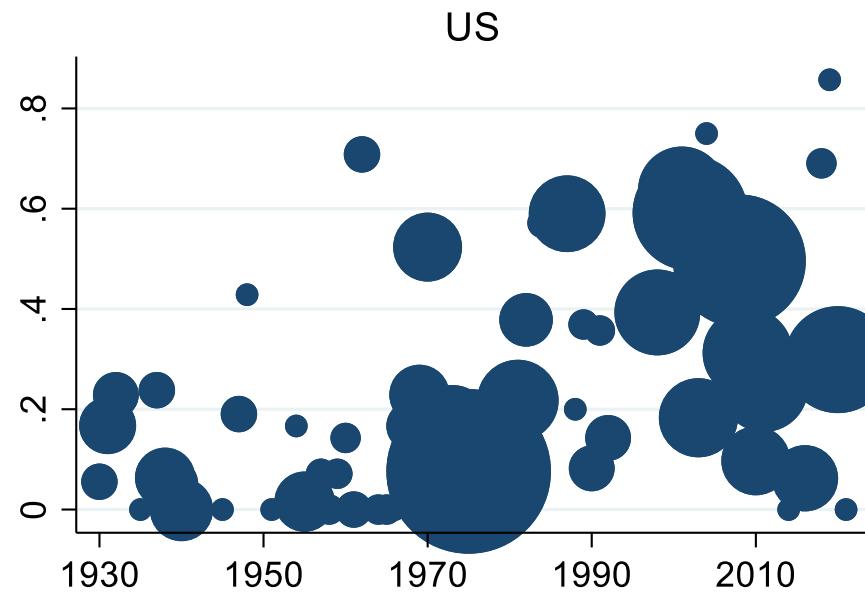
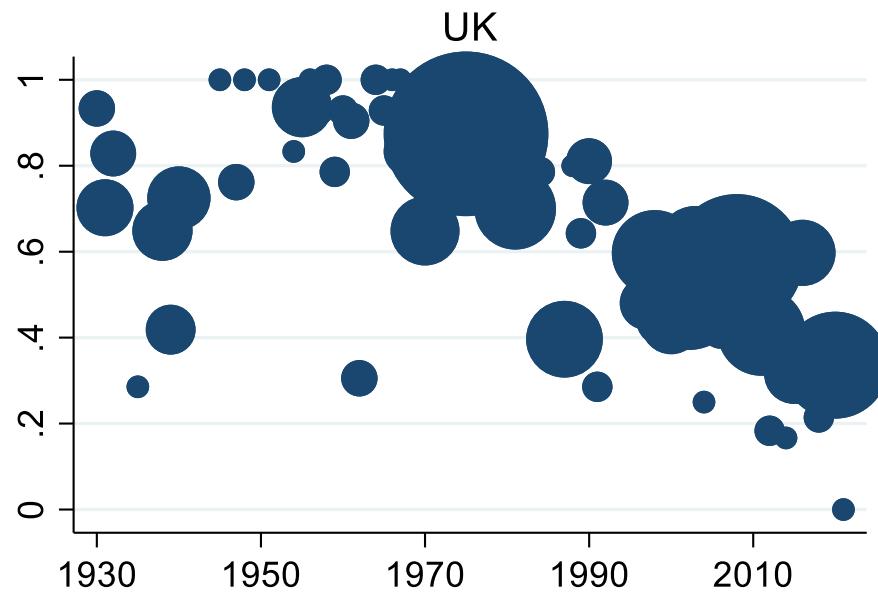
Notes: Shows average number of coders and newspaper per day, with the circle areas proportional to the number of jumps in that year. Data from 1900 to 2020. Increase in papers per day is coming from adding the Financial Times and the Houston Chronicle later in our sample.

Figure A11: UK Jumps by Year, 1930-2020



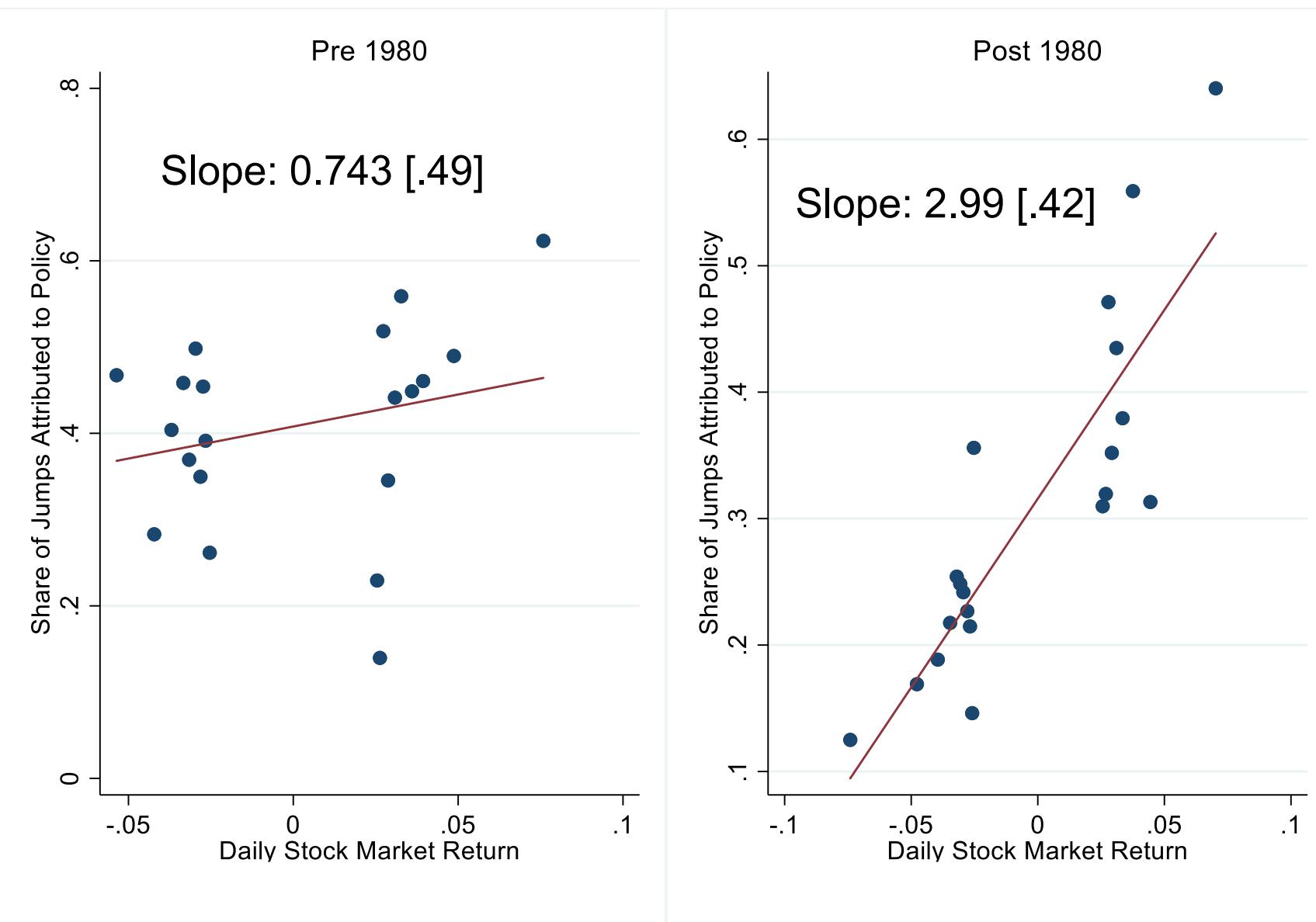
Notes: Each bar is the number of positive or negative jumps in that year. Shadings indicate the number of jumps triggered by “Policy”, “Non-Policy” and “Unknown” news. Unknown includes “no article found”. Data from 1930-2020.

Figure A12: Geographic Source of UK Jumps by Year, 1930-2020



Notes: Dot shows the share of jumps in that year in the UK by their geographic origin. The size of the dots reflects the number of jumps in that year. Data from 1930 to 2020. Excludes unknown and no article found jumps, which have no geographic attribution.

Figure A13: Policy News Also Triggers a Larger Share of Positive than Negative Jumps, Especially After 1980, in U.K. Data from 1930 to 2020

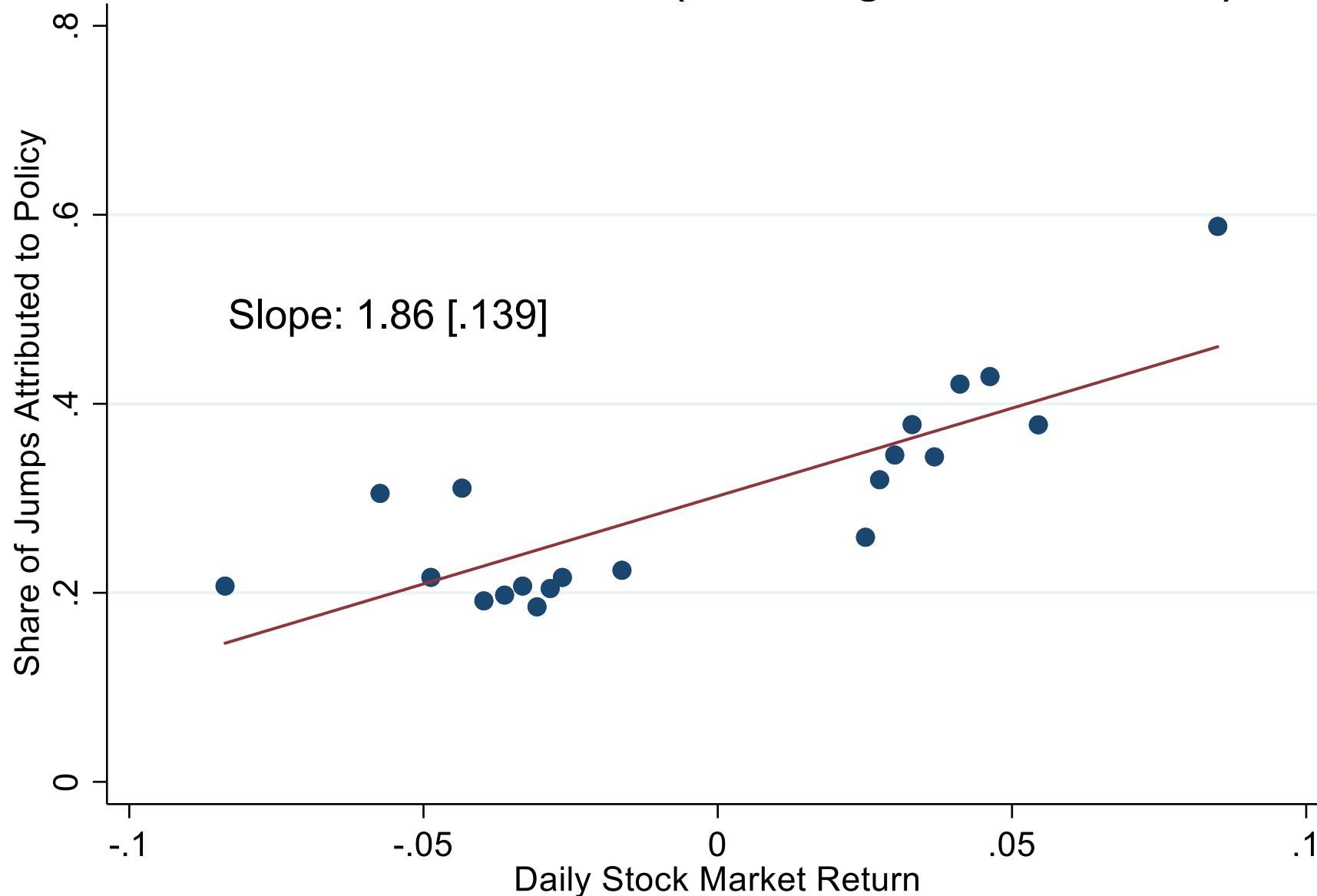


Notes: Each plot is a binscatter ($n=20$) of jump-level policy scores against jump-day stock returns, where the policy score is the fraction of the jump's codings attributed to policy-related news. For each sub-period, we regress jump-level policy scores on jump-day returns and report the t-statistic on the return variable. For jump days only, we also run the following regression,

$$\begin{aligned} \text{policy}_t \\ = a + b \text{return}_t + c 1_{\text{post}80} \\ + d \text{return}_t \times 1_{\text{post}80} + e_t \end{aligned}$$

and report the coefficient on the interaction term d , and its t-statistic at the bottom of the figure.

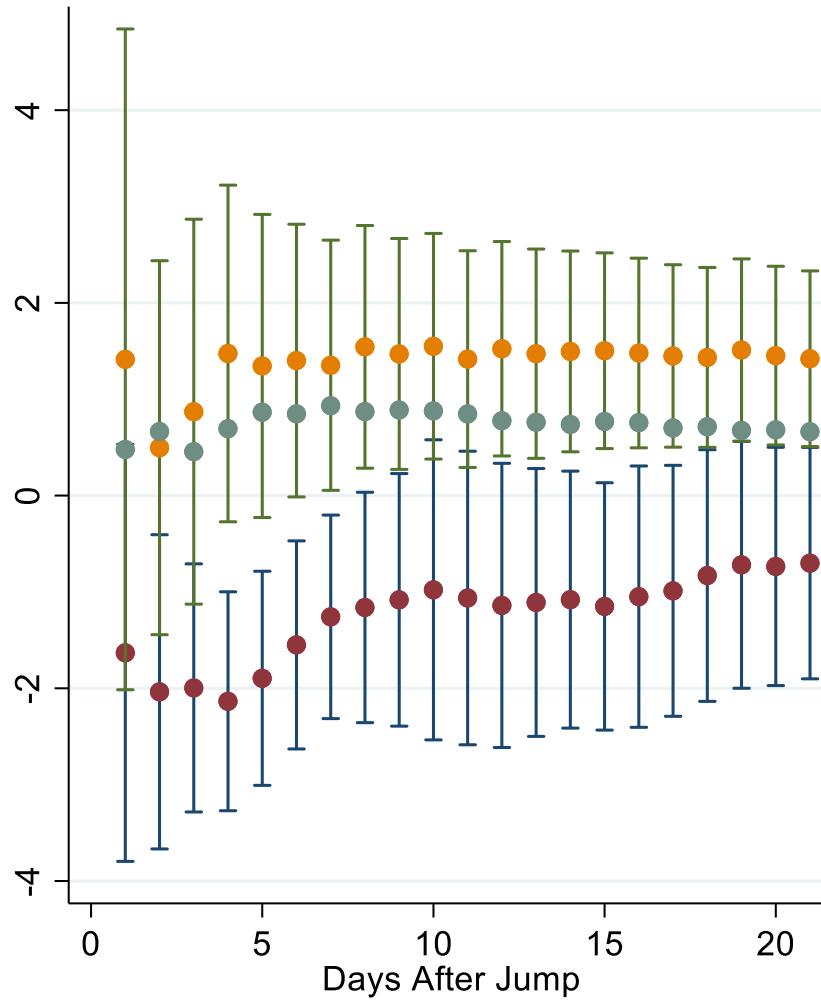
Figure A14: Policy News Also Triggers a Larger Share of Positive than Negative Jumps from 1980 to 2020 in 14 Other Countries (Excluding the U.S. and U.K.)



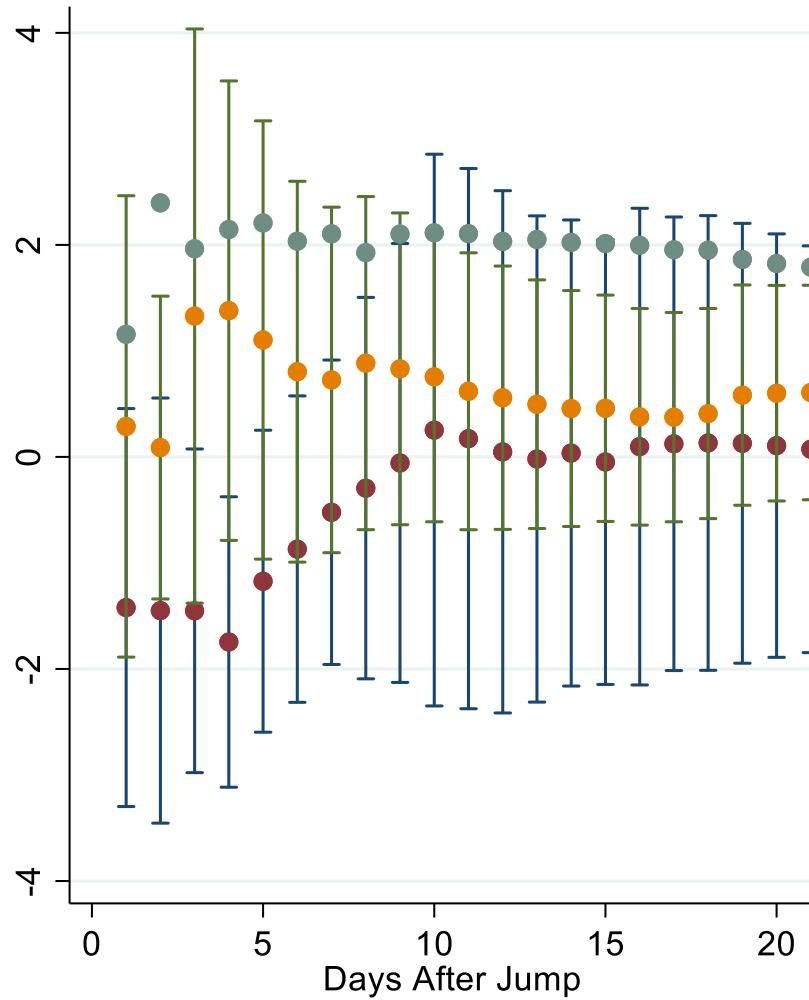
Notes: The chart shows a binscatter ($n=20$) of jump-level policy scores on jump-day stock returns from 1980 to 2020 for 14 countries (excluding the United States and the United Kingdom). The policy score is the fraction of the jump's codings attributed to policy-related news. The slope and standard error are from a regression of jump-level policy scores on a constant and jump-level returns.

Figure A15: Stock Market Volatility after Jumps Attributed to Macro News, Monetary Policy News, and All Other Types of News in U.S. Data from 1900 to 2020

All Years



Post-1945

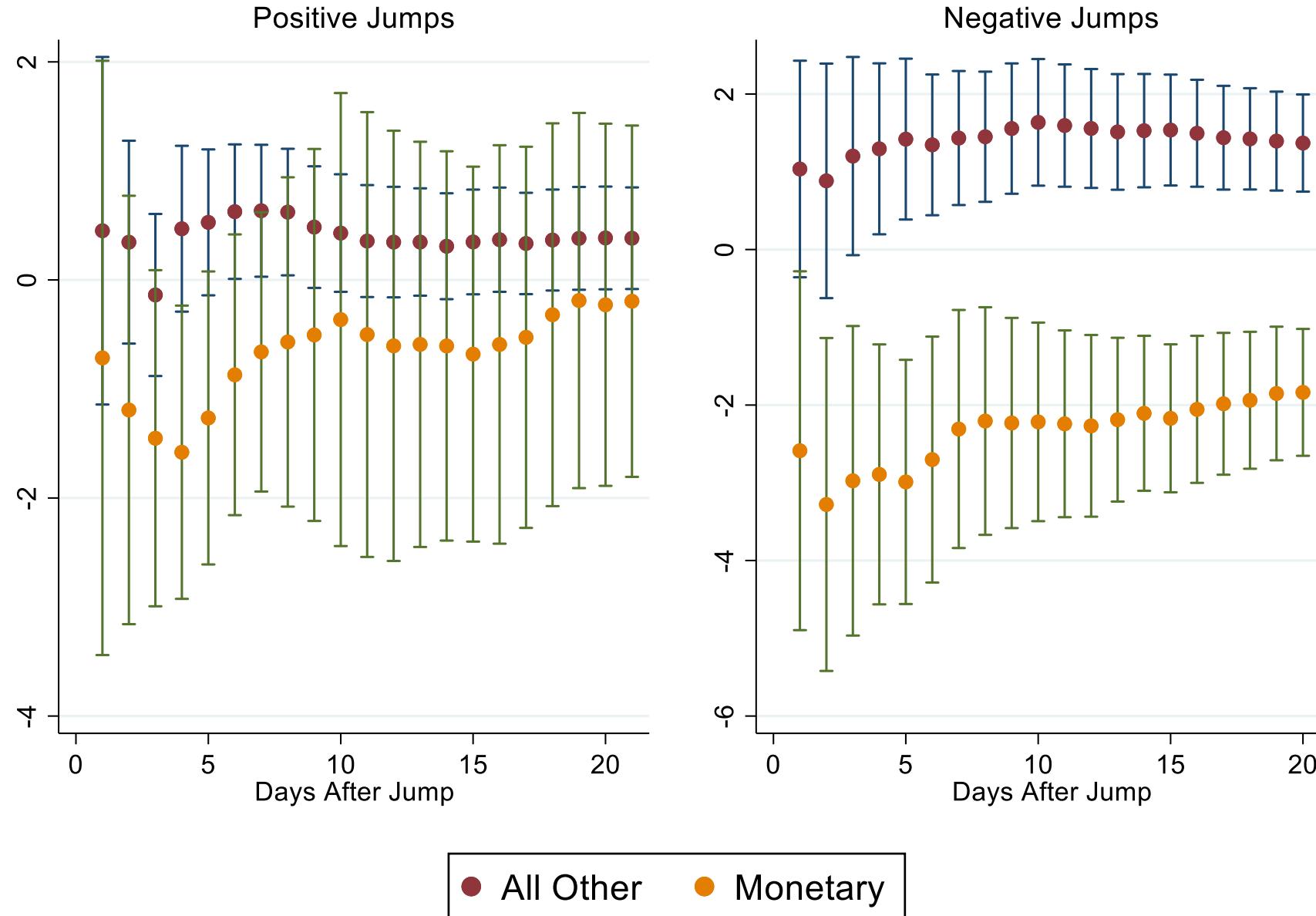


Notes: Volatility is the average squared percentage return over the next n days after a jump day. To construct the plots, we regress volatility on jump-type indicators (Monetary, Macro, All Other) and include controls for the jump-day return, split into positive and negative components, and volatility over the day, week and month prior to the jump day (HAR controls). The 95 percent confidence intervals reflect Newey-West standard errors with lags equal to 1.5 times n .



Bars represent a 95% confidence interval around the point estimate

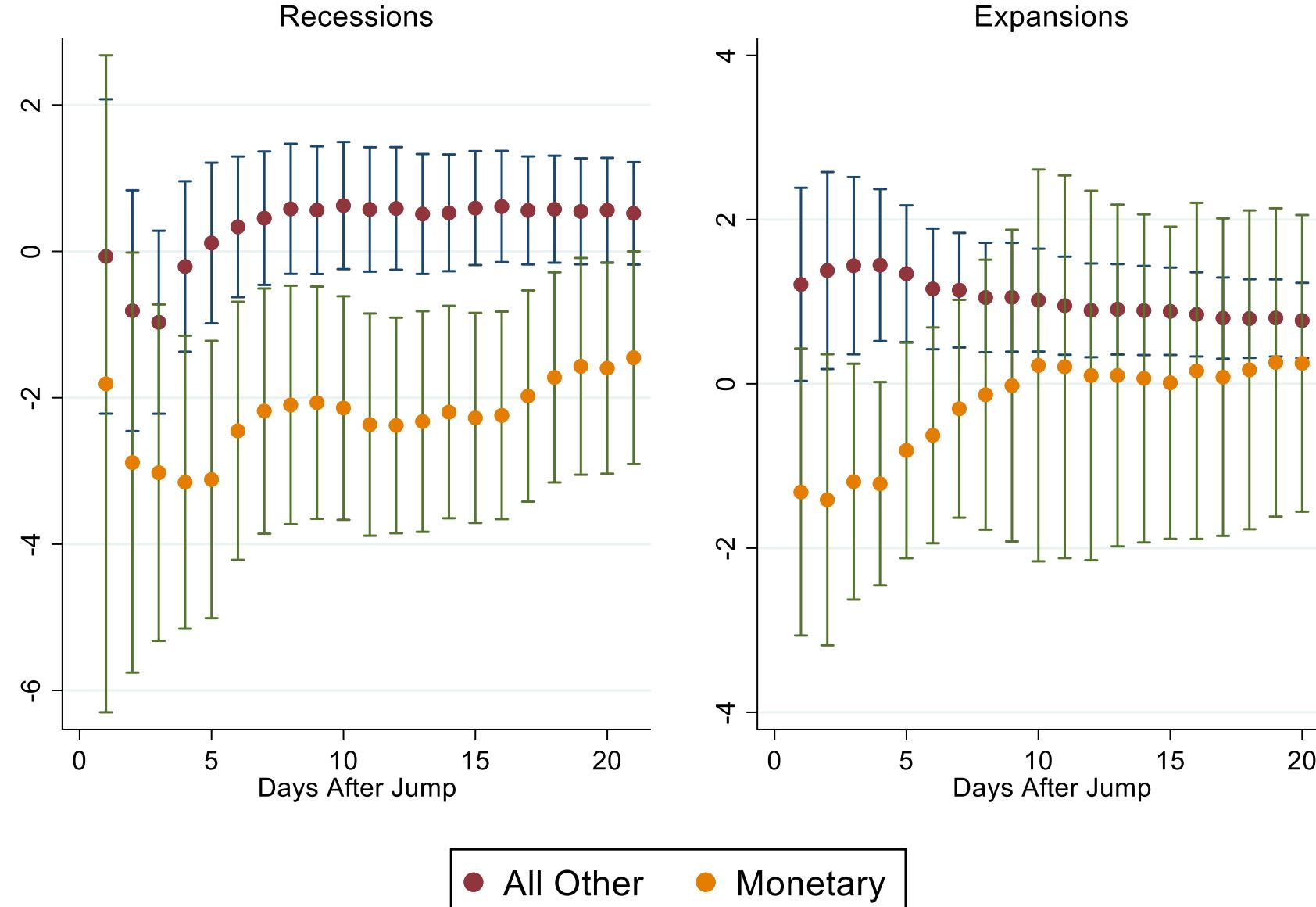
Figure A16: Volatility after Positive and Negative Jumps, Jumps Attributed to Monetary Policy and All Other News Compared, U.S. Data from 1900 to 2020



Bars represent a 95% confidence interval around the point estimate

Notes: Volatility is the average squared percentage return over the next n days after a jump day. To construct the plots, we regress volatility on four jump-type indicators (positive and negative jumps attributed to Monetary Policy, and positive and negative jumps attributed to "All Other" news) and include controls for volatility over the day, week and month prior to the jump day (HAR controls). The 95 percent confidence intervals reflect Newey-West standard errors with lags errors with lags equal to 1.5 times n .

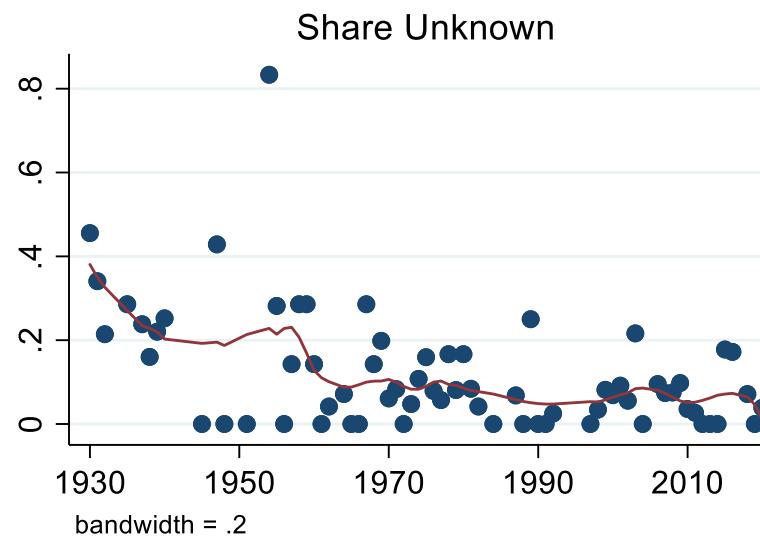
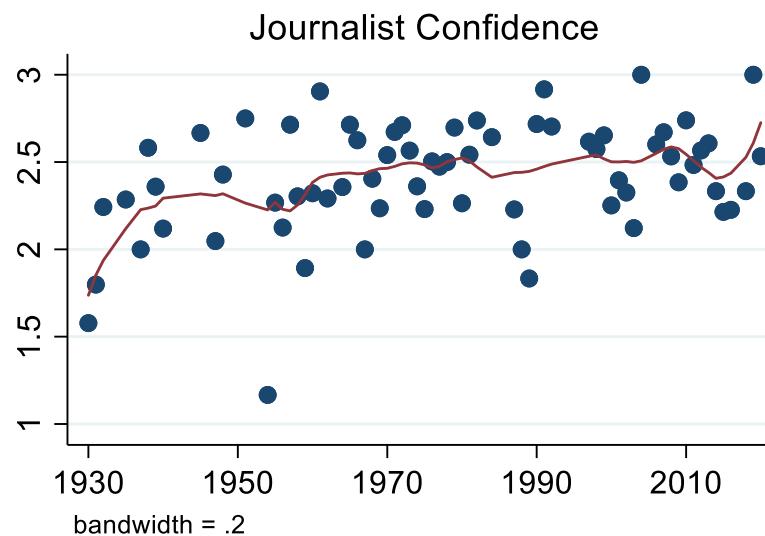
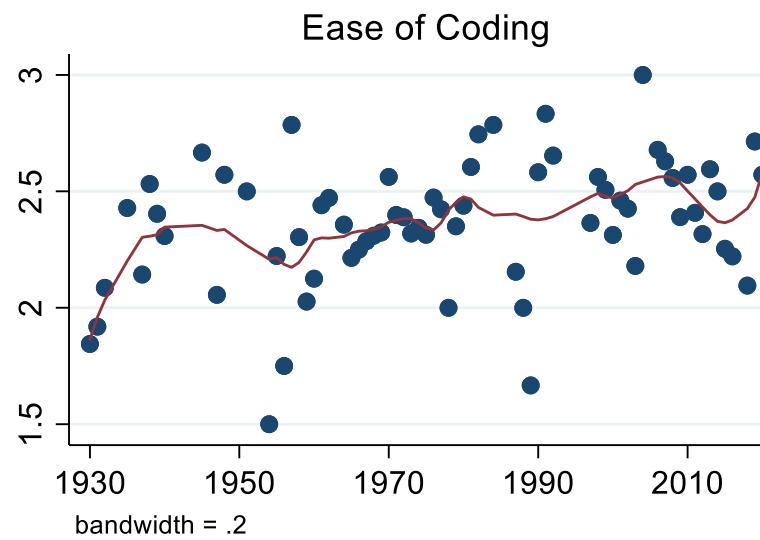
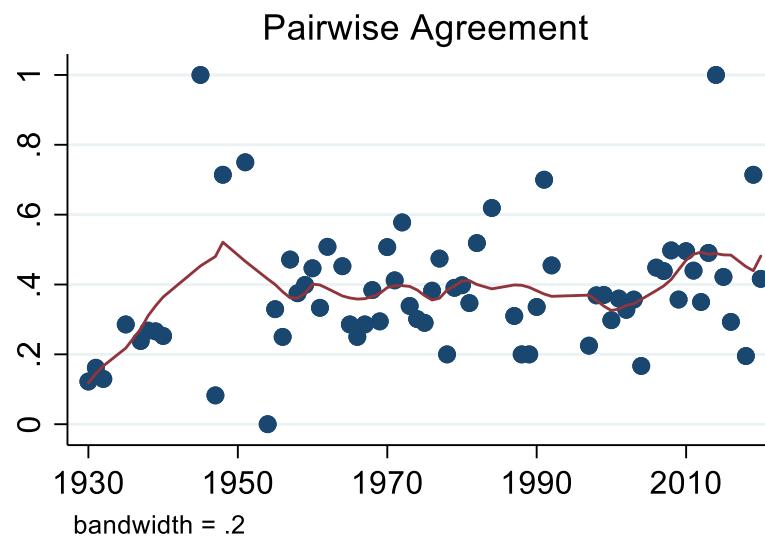
Figure A17: Volatility after Jumps during Recessions and Expansions, Jumps Attributed to Monetary Policy and All Other News Compared, U.S. Data from 1900 to 2020



Notes: Volatility is the average squared percentage return over the next n days after a jump day. To construct the plots, we split the sample of jump days into recession and expansion periods. For each subsample, we regress volatility on jump-type indicators for Monetary Policy and All Other and include controls for the jump-day return, split into positive and negative components, and volatility over the day, week and month prior to the jump day (HAR controls). The 95 percent confidence intervals reflect OLS standard errors.

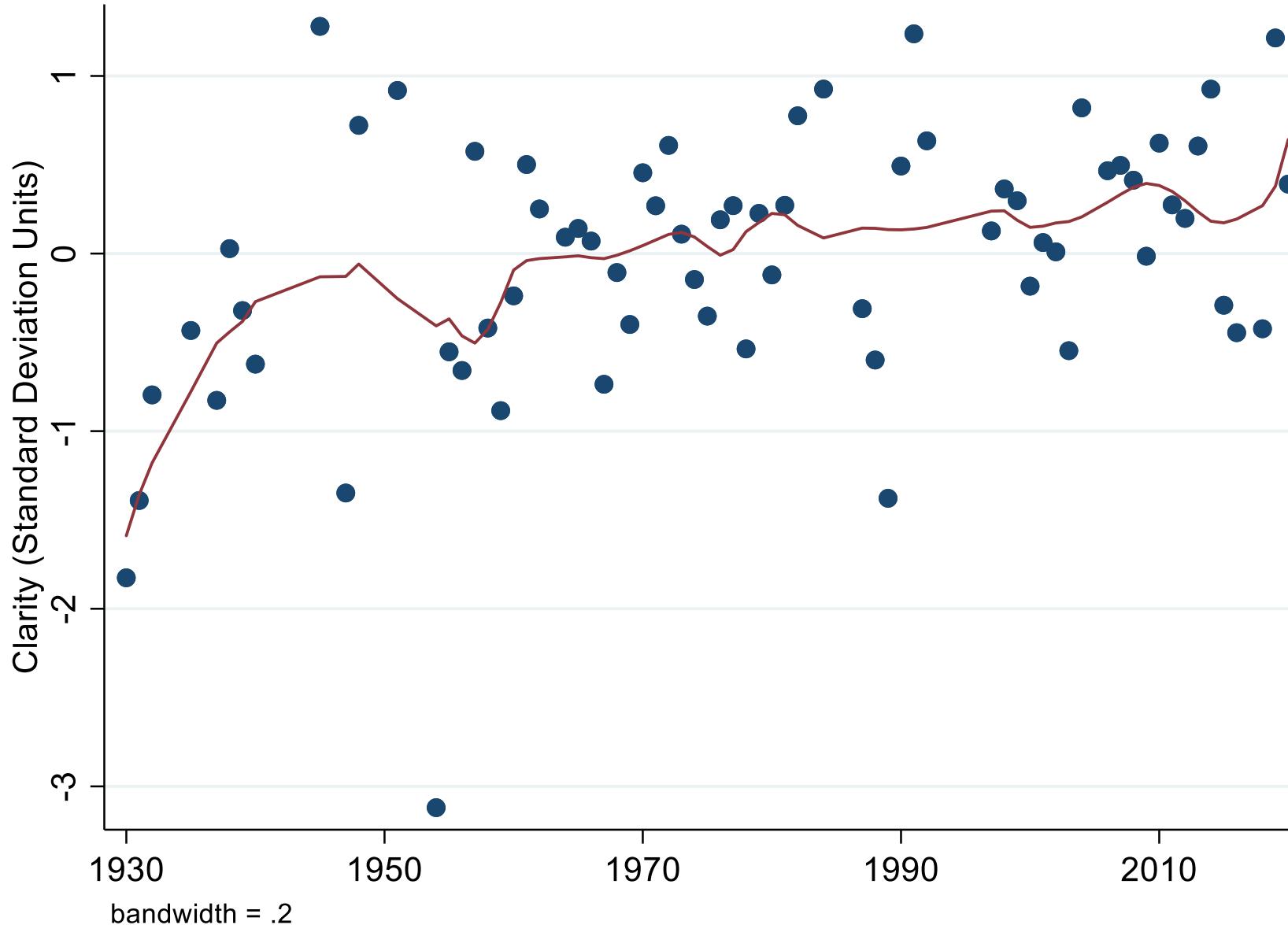
Bars represent a 95% confidence interval around the point estimate

Figure A18: Clarity Index Components Over Time, UK Data, 1930 to 2020



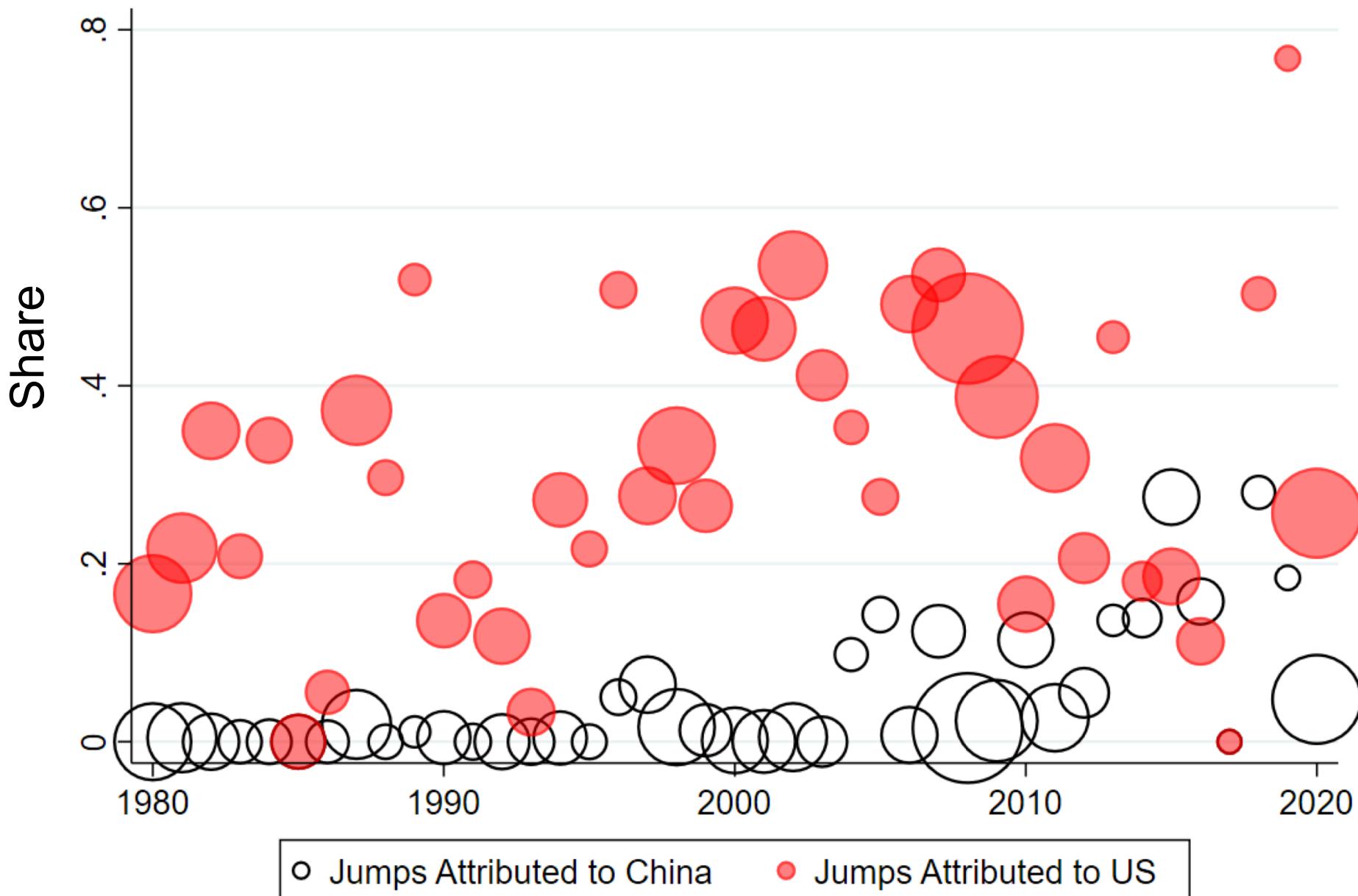
Notes: Each red line shows a LOWESS-smoothed fit to the data, with a bandwidth set to 20 percent of the data. Pairwise Agreement is the average pairwise agreement rate in the codings for a given jump. Ease of Coding is rated on a 1-3 scale, with 3 being the easiest. Journalist Confidence is rated on a 1-3 scale, with 3 being the most confident. Share Unknown is the percentage of codings for a given jump coded as "Unknown or No Explanation Offered."

Figure A19: Clarity Index Over Time, UK Data from 1930 to 2020



Notes: Clarity is the sum of Ease of Coding, Journalist Confidence, Pairwise Agreement Rate, and the share of codings attributed to “Unknown or No Explanation Offered” (multiplied by -1) after each component is scaled to zero mean and unit standard deviation one. Clarity is also scaled to have zero mean and unit standard deviation. The red line shows a LOWESS-smoothed fit to the clarity data, with a bandwidth set to 20 percent of the data. See the notes to Figure A17 for a description of each Clarity component.

Figure A20: News about the United States Triggers a Large Share of National Stock Market Jumps in Other Countries, a Pattern that Does Not Hold for Europe



Notes: This figure shows the yearly share of daily jumps attributed to the US in 15 other countries and the yearly share of daily jumps attributed to China in 14 other countries (i.e., excluding China (HK) and China (Shanghai)). The sample runs from 1980 to 2020 but does not cover all countries in all years. Dot size is proportional to the average number of jumps per country in that year. Table A1 reports the sample period by country.