What Triggers Stock Market Jumps?

Scott R. Baker[a], Nicholas Bloom[b], Steve Davis[c] and Marco Sammon[d]

November 2020

Abstract: We examine newspapers the day after major stock-market jumps to evaluate the proximate cause, geographic source, and clarity of these events from 1900 in the US, 1930 in the UK and 1980 in 12 other countries. We find four main results. First, the United States plays an outsized role in global stock markets, accounting for 35% of jumps outside the US since the 1980s. Second, policy causes a higher share of positive than negative jumps in all countries we examine, particularly monetary and government spending policy. We provide evidence that suggests these expansionary policy decisions are often made in response to poor economic conditions. Third, jumps caused by non-policy events lead to higher future volatility, while jumps caused by policy events (particularly monetary policy) reduce future volatility. Finally, the clarity of the cause of stock market jumps predicts future stock returns volatility. This type of clarity has increased substantially since 1900 as news and financial markets have become more transparent.

JEL Codes:

Keywords: uncertainty, policy uncertainty, volatility, stock market

Acknowledgements: We would like to thank the National Science Foundation and the Sloan Foundation for their financial support. We want to thanks Tasaneeya Viratyosin and Jana Obradovic for extensive research support, alongside our coder team of 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, Sophie Wang and Fernanda Kramer.

[a] Kellogg School of Management, Northwestern University and NBER
[b] Stanford University and NBER
[c] Chicago Booth School of Business and NBER
[d] Kellogg School of Management, Northwestern University

1. Introduction

A perennial question in economics and finance is "*what causes stock markets to jump*"? At one extreme is the view that all stock price movements rationally incorporate news about stock returns or discount rates. As such, large jumps in national stock indices should be accompanied by news influencing future returns or discount rates. At the opposite extreme is the view that the stock-market fluctuations are driven by speculation, for example the well-known Keynes (1936) quote that investing is like a "beauty contest", where investors price stocks not based on their opinion of their fundamental valuation but what they think others currently value them for.

In this paper we investigate this question by examining the next day's newspaper following major stock market moves. Our approach covers nearly 1,200 jumps of +/- 2.5% since 1900 in the US and over 2,500 jumps in 13 other countries. These jumps are large enough that they generally attract newspaper coverage in major newspapers the following day, so we can analyze these articles using a team of over 30 student coders. Because a sizeable fraction of aggregate returns and aggregate volatility occurs on these jump days, understanding their determinants offers insights into financial market more broadly.[1]

Our coding team categorizes stock market jumps into one of 17 categories according to the journalists' reporting, determines their geographic origin, and evaluates measures of clarity of the attributed cause. In the US, we repeat this process using five different newspapers for each jump day – the Wall Street Journal, the New York Times, the Washington Post, the Chicago Tribune and the LA Times – while in other countries we use one or two of the country's leading financial newspapers.

We also test a range of off-the-shelf machine learning and natural language models and discuss why these approaches are - at present - inferior to human codings for this task. We hope, however, that this large corpus of jump events and associated newspaper text that we develop will aid the ongoing development of text to data analysis tools and techniques for financial market moves.[2]

---

[1] Between 1900 and 2020, almost 25% of total daily variation (sum of absolute returns) and 50% of daily quadratic variation (sum of squared returns), happened on the 3% of trading days with the largest absolute returns.

[2] The jump dataset and full set of newspaper text as a training library for text-to-data algorithms is available at www.stockmarketjumps.com.

Earlier studies have examined news reports to evaluate the drivers of stock-market moves, as well. For example, classic studies by Niederhoffer (1971) and Cutler, Poterba, and Summers (1989) examined major US stock-market jumps in the past to see to what extent they could be explained by news events, coming to mixed conclusions. Our approach differs in its scale in examining around 5,000 jumps, its breadth covering 14 countries and going back to 1900 in the US (and 1930 in the UK), and the detail in measuring the causes, geographic source and clarity in a systematic way.

For some jump days, this attribution to a cause is simple. In Figure 1 plots the intraday movements (in 1-minute increments) of 4 days with daily stock market movements of greater than 2.5%. The top row contains two days with sharp, near-instantaneous, movements in the S&P 500 index which makes it easy for journalists to attribute the cause of movements on these days. In the top left panel, the market jumped over 3% after the Fed announced interest rate cuts while in the top right panel, the market plunged 2.5% at opening after an unexpectedly negative employment situation report. In the cases seen on the bottom row, the market drifted by more than 2.5% during the day, but with no clear jump or event, leaving journalists less confident about the root cause.

Leveraging these jump day categorizations and characterizations, this paper demonstrates four key results. First, the US has been and remains an extremely important driver of global stock-market volatility. Between 1980 and 2012 the share of jumps attributed to the US was 34%, substantially above its 20% share of global GDP. Moreover, this share of jumps attributed to the US has risen moderately since 1980 even as the US share of global GDP has fallen.

Second, jumps attributed to policy are disproportionately positive. In the US back to 1900 – and particularly since 1980 – as well as all other countries studied, a higher share of positive jumps were attributed to policy than negative jumps. For example, in the US since 1980 43% of positive jumps are attributed to policy while only 20% of negative jumps have a policy-related explanation. Looking at individual categories, we see monetary policy and government spending are the most likely to be positive. One explanation is that large monetary and government spending surprises tend to be expansionary in reaction to negative economic news. Another explanation we have heard anecdotally is that policy makers slowly leak market negative policy to avoid unfavorable newspaper headlines but splash market positive policy to attract attention.

Third, jumps caused by non-policy events lead to higher future realized stock-market volatility, while jumps caused by policy events (particularly monetary policy) reduce realized and implied future stock-volatility. This suggests while monetary policy surprises lead to stock-market jumps, they may reduce future volatility.

Fourth, the 'clarity' of stock market move attribution – measured by the share of articles within and across papers that agree on the cause of a jump, the share of attributions to "unknown" causes, and the confidence of the journalists' assertion over a cause - has increased significantly over our sample period. From 1900 to 1945 news coverage of financial markets shows an especially steep rise in clarity, probably linked to the improvements in financial transparency, communications and news. Clarity also turns out to matter for future volatility – perhaps unsurprisingly, jumps which have unclear attribution are followed by significantly more volatility in future days and weeks.

Our work builds on several prior literatures. First, there is a broad literature on the effects of media coverage on financial market (see e.g. Tetlock (2007)[3], Engelberg and Parsons (2011), Carlin et. al. (2014) and Manela and Moriera (2017)). We add to this literature in multiple dimensions: (1) We show that different *types* of news i.e. events attributed to different causes have different implications for future volatility (2) We show that disagreement among media coverage about the cause of a particular jump predicts higher contemporaneous and future realized volatility and (3) We highlight the importance of government policy as a driver of stock market jumps. Our results that monetary policy jumps lead to relatively lower future volatility are consistent with evidence in Pastor and Veronesi (2012), where after bad fundamental news arises, the government steps in to ameliorate the problem. Finally, we compare our human codings with classifications from off-the-shelf machine learning algorithms and are able to identify some of the pitfalls inherent to (current) automated classification of newspaper articles.

The second broad literature we contribute to is on how the clarity of financial writing affects stock returns (see e.g. Li (2008), Shiller (2017)). We contribute to this literature by constructing a new "clarity" index based on subjective human assessment of article readability, and the strength of attribution of a cause to the jump of interest. We find high clarity predicts lower volatility after the jump. We also find that jumps without a strong link to fundamental information

[3] Note that our exercise differs from Tetlock (2007) and others, in that we are interested in the ex-post attribution of stock market jumps to causes by newspapers, rather than the effect of newspaper coverage on future stock-market behavior.

on average lead to more volatility than jumps with clear connections to new economic developments.

A large literature in asset pricing has tried to quantify the share of moves in the stock market can be attributed to fundamentals like future cash flow and discount rates (see e.g. Shiller (1981), Roll (1988), Cutler, Poterba and Summers (1989)). We continue and expand upon their work, investigating what drives large stock market movements and how these causes may have important implications for the future path of asset prices and volatility.

Many papers have measured the effect of news releases on the stock market (see e.g. Birz and Lott (2011), Boudoukh et. al. (2013), Goldberg and Grisse (2013), Fernandez-Perez et. al. (2017), Fisher et. al. (2017)). We build on this in two dimensions: (1) By focusing on days with large stock market moves, there is almost always an article in the financial press offering a potential explanation (2) By having trained readers select the articles, we are more likely to be focusing on news relevant to each large move in asset markets.

Finally, much has been written about the dominance of the United States in global financial markets (see e.g. Maggiori, et. al. (2018), Boz et. al. (2017), Obstfeld (2015) and Gopinath and Stein (2018)). We contribute this literature by recording the geographic origins of the jumps in our sample and confirming the dominant role of U.S. news developments as a driver of jumps globally. In addition, Ehrmann et. al. (2011) looks at transmission of shocks both across countries and across asset classes. They find US has strong influence on Europe, but Europe has minimal effect on US. We find this is also true for stock market jumps.

Section 2 describes the construction of the categorized stock market movement data as well as the other data sources utilized in the paper and contains several exercises taken to evaluate the accuracy of the categorizations. Section 3 presents facts regarding composition of jump drivers over time and across countries, noting differential effects that jump categories have on volatility. Section 4 discusses our measurement of the clarity of jump category attribution and how this drives future stock volatility. Section 5 notes how machine learning approaches have difficulty categorizing these jump days. Section 6 concludes.

2. Data

Using a large team of human readers, we categorize the cause of large daily stock market moves based on newspaper coverage the following day. Before discussing the details, we start with four examples by reviewing the articles we used to code the jumps whose intraday price patterns are shown in the four panels of Figure 1.

The top left panel of Figure 2 displays the beginning of the Wall Street Journal article discussing the stock jump on April 18th, 2001. From the title of the article and the first sentence this clearly attributes the move to a "*surprise interest rate cut*". The geographic source of this trigger would be the US, journalist confidence would be "high" and ease of coding would be "easy". The lower right panel of Figure 2 displays the article discussing the July 2nd, 2009 jump. Again, the article makes it very clear in the title and first sentence that an "*expectedly gloomy jobs report*" led to the fall, with the geographic origin the US, journalist confidence "high" and easy of coding "easy". Overall, these are both high clarity articles as the cause of the jumps are clearly and confidently spelled out.

In Figure 3 we display two low clarity articles for 5% jump on *the same day*, December 26th, 2018. The left article came from the Wall Street Journal and would be coded as unknown since the journalist stated "*investors and traders were left scratching their heads to explain the wild swings*". Notably, the journalist wrote this three paragraphs into the article, reflecting the standard approach of placing less important or certain facts further down the article. The right article came from the New York Times and had a primary cause as consumer spending based on "*early reports of a strong holiday-shopping season*", with secondary causes of commodities and international trade also mentioned. Again, these are noted somewhat lower down the article in the second and third paragraphs.

Other newspapers on the same day gave further divergent explanations. For example, the LA Times gave three explanations running up to nine paragraphs deep into the article, including "*a late report that a U.S. government delegation will travel to China*". Overall, the jump on December 26th, 2018 was a low clarity jump according to our measures – some newspapers explicitly stated the cause of the jump was unknown, others offered multiple reasons and these differed across papers, and the papers stated these with lower confidence (deeper within the article and with less definitive language).

2.1 US Stock Jumps Data

To assemble our sample of large jump days for the United States, we first compile a list of all days where the CRSP Value-Weighted Index had an absolute return of 2.5% or more after 1926. Prior to 1926, we utilize the GFD's extension of the Dow Jones index.

In the United States, we utilize the following procedure across five major newspapers: The Wall Street Journal, the New York Times, the Chicago Tribune, the Washington Post, and the LA Times.[4] For each newspaper and each day with a market move of more than 2.5%, human readers search the newspaper's archive for relevant articles published the following day.

The readers search the archive on a given date for articles that mention phrases like 'stock market', 'wall street', 'S&P', or 'Dow Jones'. The readers select the first article that features the search terms in the title and has relevant terms in the abstract/summary of the article or mentions the previous day's percentage rise or fall in the index in the title. Readers were instructed to avoid summaries, abstracts, digests, etc. (articles <300 words). If an article satisfied these requirements but did not directly discuss the cause of the previous day's movement, additional articles were checked using the procedure define above, excluding the original article.

If none of the search terms, index terms, mentions of the rise or fall, or mentions of the previous day's market action appeared, then a more in-depth search is undertaken where several articles are read in depth and the most appropriate article chosen. With this procedure, we were able to identify at least one relevant article for every day with a large stock market move in our US post-1926 sample.[5]

Readers are assigned to carefully review each article and categorize the article's attribution of the cause of the stock market movement on the previous day. A detailed approach to this coding is laid out in the detailed (100+ page) online appendix "Coding Large Daily Financial Market Moves - Data Construction Guide".[6] For each category, a careful definition, as well as several

[4] 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.

[5] Especially in the earlier half of our sample, the most common article that is selected in the Wall Street Journal was the daily 'Abreast of the Market' column that has been utilized by other researchers for textual analysis. However, in most cases across our sample period, other articles do a more thorough job of highlighting causes of the previous day's movement.

[6] The categories are: Commodities, Corporate Earnings and Profit, Elections and Political Transitions, Foreign Stock Markets, Government Spending, Macroeconomic News, Monetary Policy and Central banking, Non-Sovereign Military/Terror, Regulation, Sovereign Military/Terror, Taxes, Trade and Exchange Rate Policy, Other Policy, Other Non-Policy, and Unknown.

examples from newspaper articles, are provided. Table 1 displays the categories that coders assign jump days to, along with the frequencies with which those categories are observed.

In addition, the Data Construction Guide goes on to further define the boundaries between pairs of related categories. As one example, the Data Construction Guide highlights that the 'Monetary Policy & Central Banking' category is distinguished from the 'Macroeconomic News & Outlook' category as follows:

> 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 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.

Each day in our sample is assigned a primary categorical cause for the day's large market movement. Many days also are coded with secondary causes, as determined by the weight put on each cause within the newspaper article. Causes that are emphasized in the title or sub-title of the article are given more weight, as are causes that are specifically noted to be the primary driver of the day's large movements. If an article mentions multiple causes but does not clearly denote a primary cause, the readers utilize the order in which the reasons are mentioned or discussed in the article as a tie breaker. Additional reasons (beyond primary and secondary) can be noted in the comment field.

For each primary cause of a market movement, the geographic source was also recorded. For instance, a large market movement in the US driven by a change in the Federal Funds Rate would be attributed to the United States, whereas a large market movement in the US caused by the decision of the UK to leave the gold standard would be attributed to the United Kingdom. Multiple countries may be cited if, for instance, a statement or action was taken by a multinational organization or coalition of countries.

Four additional measures for each article are recorded by the reader. The first is a measure of ‘Journalist Confidence’. That is, the confidence with which an article advances an explanation for a given day’s market movements. This ranges from a Confidence score of 3 (high confidence) if the article’s author directly states that the move was driven by a specific factor, to a score of 1 (low confidence) if the author gives multiple potential reasons, or directly states that investors and analysts were unsure of the reason for a market movement.

Readers also classify articles based on the ‘Ease of Coding’, which measures how difficult it was to assign a primary cause to the market movement. The score ranges from 3 (Easy to code) for articles that rapidly and clearly identify the cause of the jumps to 1 (Hard to code) for articles that meander, offer several explanations or are hard to understand. This measure is correlated with Confidence but is not the same – a journalist may be confident that specific events drove markets on a given day but write an opaque article, or may lay out a complex cause that touches on several of our categories at once.

Related to these two measures, the coders record how far into the article they had to read before they were confident about the primary category for the jump. The score ranges from 3 (Clear from Title) for articles that advance a clear explanation in the title to 1 (Not Clear Until Reading Beyond the First Two Paragraphs) for articles that have uninformative titles, and take several paragraphs to get to the main explanation. Regardless of how this field is coded, readers were instructed to always review the whole article to be sure that the key information is where they said it was.

Finally, coders recorded a key passage. This field is the passage in the article that was most important in determining their coding. This differs from the detailed cause/notes field, in which the coders are paraphrasing the article. An article with a primary category of Taxes might have a key passage like, “The completion of a tax deal between the White House and Congress sent stocks soaring Wednesday.”

For the United States, we conducted a thorough cross-validation with an average of 8.9 coders across multiple newspapers for each day.[7] 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 reader 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, readers still sometimes disagreed regarding some moves that were highly uncertain. For such days, readers could 'agree to disagree' regarding the causes of the stock move and our final dataset reflects such persistent disagreement.

Finally, before analysts started coding, they carefully read the coding guide, underwent a half-day training session and then coded 50 WSJ training articles over the next two to 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 1900-2020H1, 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

[7] See Appendix Figure A5 for a timeline of the number of coders who read and coded articles for each day's jump. The number of newspapers per day increases later in the sample, as we added the US edition of the Financial Times and the Houston Chronicle.

countries with more volatile stock-markets we increased the, choosing these thresholds to cover approximately 2-3% of trading days.[8]

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. Readers 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). 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.

2.3 Validation of Human Coder Data

A potential concern is the reliability of human readers in consistently identifying the correct 'category' of the cause for a given large stock market movement. 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 41% agree on the 17 granular categories. While this may not seem high, if we randomly assign coders to categories from the unconditional

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

distribution in Table 1, agreement would be only 12%. 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 53% agreement rate on the granular categories.

Agreement also increases when considering only pairs of coders reading the same newspaper.[9] 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 over 20%. An increase in the reported ease of coding has an effect of a similar magnitude.

Among papers, agreement is highest for readers 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 over 90% agreement rate on the policy vs. non-policy split, and an almost 80% agreement rate on the granular categories.

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 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 develop another validation exercise

[9] 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.

to assess the accuracy of our newspaper-based classifications of jump reasons based on relative industry returns on jump days. The results of that exercise also support the view that our newspaper-based explanations are informative. See Table A4 and the related discussion.

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 2020H1. 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).[10]

In the figure, we see three particularly notable spikes in the frequency of jumps: the first starting during the Great Depression from the late 1920s until the late 1930s; the second during the Great Recession from 2008-2012, and the third the COVID-19 pandemic. While the years since the waning of the Great Recession (2010-2019) had seen only 5.8 jumps per year on average in the United States, there were 33 large jumps in just the first six months of 2020, where our data ends. In fact, the month from February 24th to Mach 24th contained 18 market jumps across 22 trading days, more than any other one month period in our sample back to 1900.

There were also several periods of higher volatility during the early 1900s, with World War I, the Panic of 1907, and other financial panics playing a role. Perhaps surprisingly, other wars like World War II, the Korean War, and the War in Vietnam did not produce many large daily jumps in the stock market. During the post-war period, there are long periods with few daily movements large enough to cross the threshold of our sample.

Figure B1 mirrors Figure 4 using data regarding stock market jumps in the United Kingdom back to 1930. Here we find a strikingly different pattern of jump days, with little stock market

[10] 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 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.

volatility during the Great Depression years and a large surge in the early 1970s during a severe recession and IMF crisis.

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 1930-2020H1 and those from our sample of 12 foreign countries[11]. This shows that not only have 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 were driving 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.

From the table we take away two important stylized facts. First, policy news drives a large portion of large daily stock market movements. Over 37% of US jumps are attributed to policy: more than macroeconomic news (24%) and corporate earnings (11%) combined. Globally, 26% 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 second largest overall mover, reflecting in particularly the role of US stock markets in driving global market movements. Third, there is a surprisingly high number of jumps for which the newspapers reported the driver explicitly as unknown, totaling 15.6% and 11.3% in the US and ROTW respectively.[12]

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 10-year Treasury Bills of more than 15 basis

---

[11] Australia, Canada, China, Germany, Greece, Ireland, Japan, New Zealand, Saudi Arabia, Singapore, South Africa, and South Korea. We utilize two separate sets of observations from China, one from the Hong Kong stock exchange and one from the Shanghai stock exchange as these indexes cover different portions of the Chinese economy.

[12] In Appendix Table A2 we look at the major movers in the US market by decade. Consistent with Table 1, we see that before 1945 sovereign military was one of the major drivers due to World War I and II, while post war macro news, corporate earnings and monetary policy dominates.

points. Here we find a significantly different distribution of categories than with equity jumps, both in the United States and internationally. Jumps in bond markets tend to be driven almost entirely by macroeconomic news and news about monetary policy, which collectively account for 80% of bond market jumps.

3.3 Geographical Source of Stock Market Jumps

Going beyond the categorical cause of large stock moves, we examine the geographic sources of large jump. Figure 5 plots timelines of the geographic source of large stock market jumps in the US, showing the dominance of US news for US jumps. On average, US jumps are attributed to a cause at least partially driven by domestic events 87% of the time. The only periods in which the US share persistently falls below 50% are during the First and Second World Wars (when Europe and Asia become important drivers) and during the European debt crisis in the early 2010s. Figure B2 of the Appendix presents the same plot for the UK: Consistent with the growing global dominance of US financial markets, the UK's share of UK-sourced jumps declined from 70% between the 1930s and the 1960s to around 25% from 2010 onwards. We also see on the top right Europe's contribution to US stock market volatility appears to be falling while Asia's has been recently rising 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 the US and Europe back to 1980 in third party countries – Australia, Canada, China (HK), China (Shanghai), Japan, New Zealand, Saudi Arabia, Singapore, South Africa, and South Korea. On the same plot we report the global share of US and European GDP. One striking finding is that while Europe has a slightly greater share of global GDP over this period its share of attribution for jumps in third party countries is 4.5%, almost a fifth of the US's 27% share. And although the US global share of GDP has been slowly declining, the share of jumps attributed to the US has also been increasing over this period (with the time-trend significant at the 5% level).

3.4 Differences in 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.[13] 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 steepening over time - also hold in the UK (Appendix Figure B6). 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 A5).

Examining the individual categories (Appendix Table A6) we see that monetary policy and government spending jobs 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 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. 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. For each of the three panels, the x-axis represents the cumulative return in the CRSP Value-Weighted index over the past quarter (66 trading days). In the top left panel, we plot a bin-scatter of the share of jumps attributed to policy against the returns over the past quarter, but only for *positive jumps*. We find a negative slope, which implies that positive stock market jumps are more likely to be driven by policy after bad returns. The top

[13] Figure A7 replicates this figure for all the Non-US and Non-UK countries in our sample revealing a similar result.

right panel is identical, but only includes *negative jumps*. Here, we find a positive slope, which implies that negative stock-market jumps are more likely to be driven by policy following good returns. Finally, in the bottom left panel we construct a 'net' policy score, which takes a value of 1 if the jump is positive and policy, 0 if it is non-policy, and -1 if the jump is negative and attributed to policy. 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.

3.5 Differences in Volatility by Stock Market Jump Category

We have documented the fact that the categorical causes and geographic origins of stock market movements vary across countries and have changed substantially over time. We now turn to whether these categorical differences in the cause of large stock market movements can predict future differences in financial market outcomes.

We find that, for a given size of stock market move, the reasons behind the move have systematically different implications for realized market volatility in the following days and weeks. We measure realized volatility over an *n* day horizon as the mean squared return on the CRSP Value-Weighted index over those *n* days. We use the uncentered second moment to avoid the difficulties inherent in measuring the mean stock return over a short horizon.

While all jump days lead to elevated levels of volatility, we test whether some types of jumps have more persistent effects than others, utilizing the following regression approach:

$$\mathbf{100}\sum_{i=1}^{n}\frac{r_{t+i}^{2}}{n} = a + b\left(r_t \times \mathbf{1}_{r_t>0}\right) + c\left(|r_t| \times \mathbf{1}_{r_t\leq 0}\right) +$$

$$d\left(r_{t-1}^{2}\right) + e\left(\sum_{i=1}^{5} r_{t-i}^{2}\right) + f\left(\sum_{i=1}^{22} r_{t-i}^{2}\right) +$$

$$+ g\,monetary_t + h\,allother_t + Fixed\ Effects + e_t$$

$\boldsymbol{r_t}$ is the return on the CRSP value-weighted index. The left-hand side term is the average realized volatility over an n-day horizon. The first set of right-hand side variables are controls for the day's return, and allowing for an asymmetric effect of positive and negative returns on future volatility (see e.g. Black (1976)). The second set of RHS variables are Heterogeneous Autoregressive (HAR) controls to account for the effect of persistent volatility over different horizons (Corsi 2009). The last set of RHS variables represents our jump categories. For example, $\boldsymbol{monetary_t}$ will take the value 1 if all coders reading the article on that day classified the article as Monetary Policy, and will take the value 0 if no coders assigned the article Monetary Policy. For days with disagreement between coders as to the primary category, the variable will take a value between zero and one.[14] This regression includes all days (including non-jump days), and $\boldsymbol{monetary_t}$ will take a value of 0 on any non-jump day as well.

We find strong evidence that large stock jumps driven by monetary policy produce realized volatility substantially lower than those driven by all other categories. We plot coefficients from this regression in Figure 9, looking at the 22 trading days (an approximate one-month period) after a jump day. We hypothesize that some of these differences are driven by the fact that some types of events, such as a bad macro unemployment report, may *generate* uncertainty while others, such as monetary policy announcements about a rate change, may *resolve* uncertainty. These differences are economically significant, with volatility being almost 1 standard deviation lower 10 days after a monetary jump than after a jump attributed to all other categories.

The results in Figures 9 also seem robust to a number of other cuts. In a series of appendix figures (A1, A2 and A3) we break out macro jumps, separate positive and negative jumps, and split by recessions and expansions. In all cases we find monetary policy jumps are followed by lower volatility over the next 22 trading days.

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":

[14] Fixed effects include decade indicator variables, as well as a NBER recession indicator variable, though results are robust to year fixed effects or a year trend instead of decade dummies.

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 can also combine these into a 'clarity index' by taking the first principal component factor.[15] 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 in Appendix Figures B3 and B4.

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

[15] Our results are robust to using a z-score which takes the average of each component after normalizing it to a mean of 0 and a standard deviation of 1.

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.[16] 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 sample design in 1964 and incremental improvements over the next 25 years or so. Annual CES benchmarking to universe-level employment data began in 1982. After much criticism and, occasionally, very large benchmark revisions, the BLS 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 160,000 in 1975 and 425,000 by 1989. As of 2016, the CES surveys about 620,000 business and government worksites each month. The BLS also first issued seasonally adjusted CES statistics in 1954.

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 were fairly ambiguous. Intriguingly, clarity has also appear to have fallen somewhat 2016 under the Trump administration.

[16] This paragraph draws on Johnson (2016), Kelter (2016) and Mullins (2016).

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 5-minute window divided by the total movement during the day – on our clarity index in Table 4. For each day, we calculate the absolute returns in 5-minute intervals, with the first window being 9:30AM to 9:35AM and the final window 3:55PM – 4:00PM. We then divide the largest absolute move by the sum of all the absolute moves to obtain our concentration measure.

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 0.0836 and a standard deviation of 0.05513, this result implies that a two-standard deviation shock to clarity is associated with an increase in concentration of 0.0487, or a 0.87 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 4 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 before and 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, presumably because markets movements are noisier. 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.[17] Indeed, this suggests that one reason for the persistence of volatility in financial markets is 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 there has been an ongoing debate over the role of policy in driving economic uncertainty (e.g. Baker et al. 2016) our results suggest that policy driven jumps tend to have higher clarity and induce less future volatility that non-policy driven jumps.

In Table 5, we regress future stock market volatility – in particular the squared returns over the next five days after each jump on our clarity index. 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 10.1, or a 0.2 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.

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

[17] 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.

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.[18]

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.[19] The composition of journal articles has also changed significantly over the past

[18] 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.

[19] 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.

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.

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.[20] The full details of this approach are laid out in Appendix A2.

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 typically 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).[21] 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.

---

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

[21] We display these results graphically in Appendix Figure A6.

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.

6. Conclusion

We examine newspapers the day after major stock-market jumps to catalog the proximate cause, geographic source, and clarity of these events from 1900 in the US and 1980 (or earlier) in 13 other countries. We find four main results. First, the United States plays an outsized role in global stock-markets, accounting for 35% of jumps outside the US since 1980s, far above its 15-20% share of GDP. This matches other evidence on the dominance of the US in global finance. Second, policy causes a significant higher share of positive than negative stock market jumps, both in the US and all other 13 countries we examined. Monetary policy and government spending jumps are the most strongly over-represented in positive jumps, suggesting major policy announcements are usually in respond to negative economic shocks.

Third, jumps caused by non-policy events lead to higher future stock-volatility, while jumps caused by policy events, monetary policy in particular, reduce future stock-volatility. This suggests while monetary policy surprises lead to stock-market jumps, they may reduce future volatility. Finally, the clarity of the cause of stock market jumps has been increasing since 1900, presumably because news and financial markets has become more transparent. This clarity tends to have consequences for financial markets, with higher levels of clarity predicting significantly lower volatility.

References

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold & Clara Vega, 2007. "Real-Time Price Discovery in Global Stock, Bond and Foreign Exchange Markets," *Journal of International Economics,* 251-277.

Baker, Scott R., Nicholas Bloom, Steven J. Davis and Marco Sammon, 2019. "Coding Large Daily Financial Market Moves: Data Construction Guide," 24 August, at https://stockmarketjumps.com/files/coding_guide_pdf.pdf.

Baker, Scott R., Nicholas Bloom, Steven J. Davis, Kyle Kost and Marco Sammon, and Tasaneeya Viratyosin, 2020 "The Unprecedented Stock Market Reaction to COVID-19" University of Chicago Working paper.

Barro, Robert J. "The stock market and investment." The Review of Financial Studies 3.1 (1990): 115-131.

Birz, Gene, and John R. Lott Jr. "The effect of macroeconomic news on stock returns: New evidence from newspaper coverage." *Journal of Banking & Finance* 35.11 (2011): 2791-2800.

Black, Fischer, "The Dividend Puzzle". The Journal of Portfolio Management Winter 1976, 2 (2) 5-8

Blei, David, Ng, Andrew, and Jordan, Michael. "Latent Dirichlet allocation". The Journal of Machine Learning Research, Volume 3, 3/1/2003; Pages 993-1022

Bollerslev, Tim. "Generalized autoregressive conditional heteroskedasticity." Journal of econometrics 31.3 (1986): 307-327.

Bollerslev, Tim, Robert F. Engle, and Daniel B. Nelson. "ARCH models." Handbook of econometrics 4 (1994): 2959-3038.

Boudoukh, Jacob, et al. *Which news moves stock prices? A textual analysis*. No. w18725. National Bureau of Economic Research, 2013.

Boz, Emine, Gita Gopinath, and Mikkel Plagborg-Møller. *Global trade and the dollar*. No. w23988. National Bureau of Economic Research, 2017.

Brogaard, Jonathan, and Andrew Detzel. "The asset-pricing implications of government economic policy uncertainty." *Management Science* 61.1 (2015): 3-18.

Campbell, John Y., et al. "Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk." The Journal of Finance 56.1 (2001): 1-43.

Carlin, Bruce I., Longstaff, Francis A. and Matoba, Kyle, (2014), Disagreement and asset prices, Journal of Financial Economics, 114, issue 2, p. 226-238.

Corsi, Fulvio "A Simple Approximate Long-Memory Model of Realized Volatility" Journal of Financial Econometrics, Volume 7, Issue 2, 1 March 2009, Pages 174–196

Croux, Christophe, and Peter Reusens. "Do stock prices contain predictive power for the future economic activity? A Granger causality analysis in the frequency domain." *Journal of Macroeconomics* 35 (2013): 93-103.

Cutler, David M., James M. Poterba and Lawrence H. Summers. "What Moves Stock Prices?" The Journal of Portfolio Management Spring 1989, 15 (3) 4-12

Dougal, Casey, et al. "Journalists and the stock market." The Review of Financial Studies (2012): 639-679.

Ehrmann, Michael, Marcel Fratzscher, & Roberto Rigobon. "Stocks, bonds, money markets & exchange rates: measuring international financial transmission." Journal of Applied Econometrics (2011): 948-74.

Engelberg, Joseph E., and Christopher A. Parsons. "The causal impact of media in financial markets." The Journal of Finance 66.1 (2011): 67-97.

Engle, Robert F. "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation." Econometrica: Journal of the Econometric Society (1982): 987-1007.

Fama, Eugene F. "Stock returns, real activity, inflation, and money." The American economic review 71.4 (1981): 545-565.

Fischer, Stanley, and Robert C. Merton. "Macroeconomics and finance: The role of the stock market." (1984).

Fernandez-Perez, Adrian, Bart Frijns, and Alireza Tourani-Rad. "When no news is good news–The decrease in investor fear after the FOMC announcement." *Journal of Empirical Finance* 41 (2017): 187-199.

Fisher, Adlai J., Charles Martineau, and Jinfei Sheng. "Media attention, macroeconomic fundamentals, and the stock market." *University of British Columbia Working paper* (2017).
Flannery, Mark J. and Aris A. Protopapadakis, 2002. "Macroeconomic Factors *Do* Influence Aggregate Stock Returns," *Review of Financial Studies,* 15, no. 3 (April), 751-782.
Glosten, Lawrence R., Ravi Jagannathan, & David E. Runkle. "On the relation between the expected value and the volatility of the nominal excess return on stocks." The journal of finance 48.5 (1993): 1779-1801.
Goldberg, Linda S., and Christian Grisse. *Time variation in asset price responses to macro announcements*. No. w19523. National Bureau of Economic Research, 2013.
Gopinath, Gita, and Jeremy C. Stein. *Banking, Trade, and the making of a Dominant Currency*. No. w24485. National Bureau of Economic Research, 2018.
Guo, Hui. "Stock market returns, volatility, and future output." *Review-Federal Reserve Bank of Saint Louis* 84.5 (2002): 75-84.
Gurkaynak, Refet; Brian Sack; Eric Swanson, 2005. "Do Actions Speak Louder than Words? The Response of Asset Prices to Monetary Policy Actions and Statements," Macroeconomics 0504013, University Library of Munich, Germany.
Johnson, Nicholas, 2016. "One Hundred Years of Current Employment Statistics Data Collection," *Monthly Labor Review,* U.S. Bureau of Labor Statistics, August.
Kelly, Bryan, Lubos Pastor, and Pietro Veronesi "The price of political uncertainty: Theory and evidence from the option market". Journal of Finance 71, 2417-2480, 2016.
Kelter, Laura A., 2016. "One Hundred Years of Current Employment Statistics – The History of CES Sample Design," *Monthly Labor Review,* U.S. Bureau of Labor Statistics, August.
Keynes, John Maynard, 1883-1946. The General Theory of Employment, Interest and Money. London: Macmillan, 1936. Print.
Li, Feng. "Annual report readability, current earnings, and earnings persistence." *Journal of Accounting and economics*45.2-3 (2008): 221-247.
Lucca, David O., and Emanuel Moench. "The pre-FOMC announcement drift." *The Journal of Finance* 70.1 (2015): 329-371.
Maggiori, Matteo, Brent Neiman, and Jesse Schreger. *International currencies and capital allocation*. No. w24673. National Bureau of Economic Research, 2018.
Manela, Asaf, and Alan Moreira. "News implied volatility and disaster concerns." *Journal of Financial Economics* 123.1 (2017): 137-162.
Mehl, Arnaud. "Large global volatility shocks, equity markets and globalization: 1885-2011." (2013).
Muir and Moriera (2017) "Volatility-Managed Portfolios". Journal of Finance. Vol 72, Issue 4, Pg 1611-1644
Mullins, John P., 2016. "One Hundred Years of Current Employment Statistics – An Overview of Survey Advancements," *Monthly Labor Review,* U.S. Bureau of Labor Statistics, August.
Niederhoffer, Victor. The Analysis of World Events and Stock Prices. The Journal of Business, 1971, vol. 44, issue 2, 193-219. Date: 1971
Obstfeld, Maurice. "Trilemmas and trade-offs: living with financial globalization." (2015).
Pastor, Lubos, and Pietro Veronesi. "Uncertainty about government policy and stock prices." *The Journal of Finance* 67.4 (2012): 1219-1264.
Pastor, Lubos & Veronesi, Pietro, 2013. "Political uncertainty and risk premia," Journal of Financial Economics, Elsevier, vol. 110(3), pages 520-545.
Roll, Richard "R2" Journal of Finance, Volume 43, No. 2, July 1988
Russell, Stuart, Peter Norvig, "A Modern Approach" Artificial Intelligence. Prentice-Hall, Englewood Cliffs 25 (27), 79-80
Shiller, Robert, (1981), "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?" American Economic Review, 71, issue 3, p. 421-36.
Shiller, Robert J. "Narrative economics." *American Economic Review* 107.4 (2017): 967-1004.
Tetlock, Paul C. "Giving content to investor sentiment: The role of media in the stock market." *The Journal of Finance* 62.3 (2007): 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 readers – 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:

Tri_{it} = 1, if the detailed description for t implies an amplified response of R_{it};

= -1, if the detailed description for t implies a dampened response of R_{it};

= 0, otherwise.

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.[22]

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 A4. 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.

[22] 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 A4 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 A4 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 A4 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.

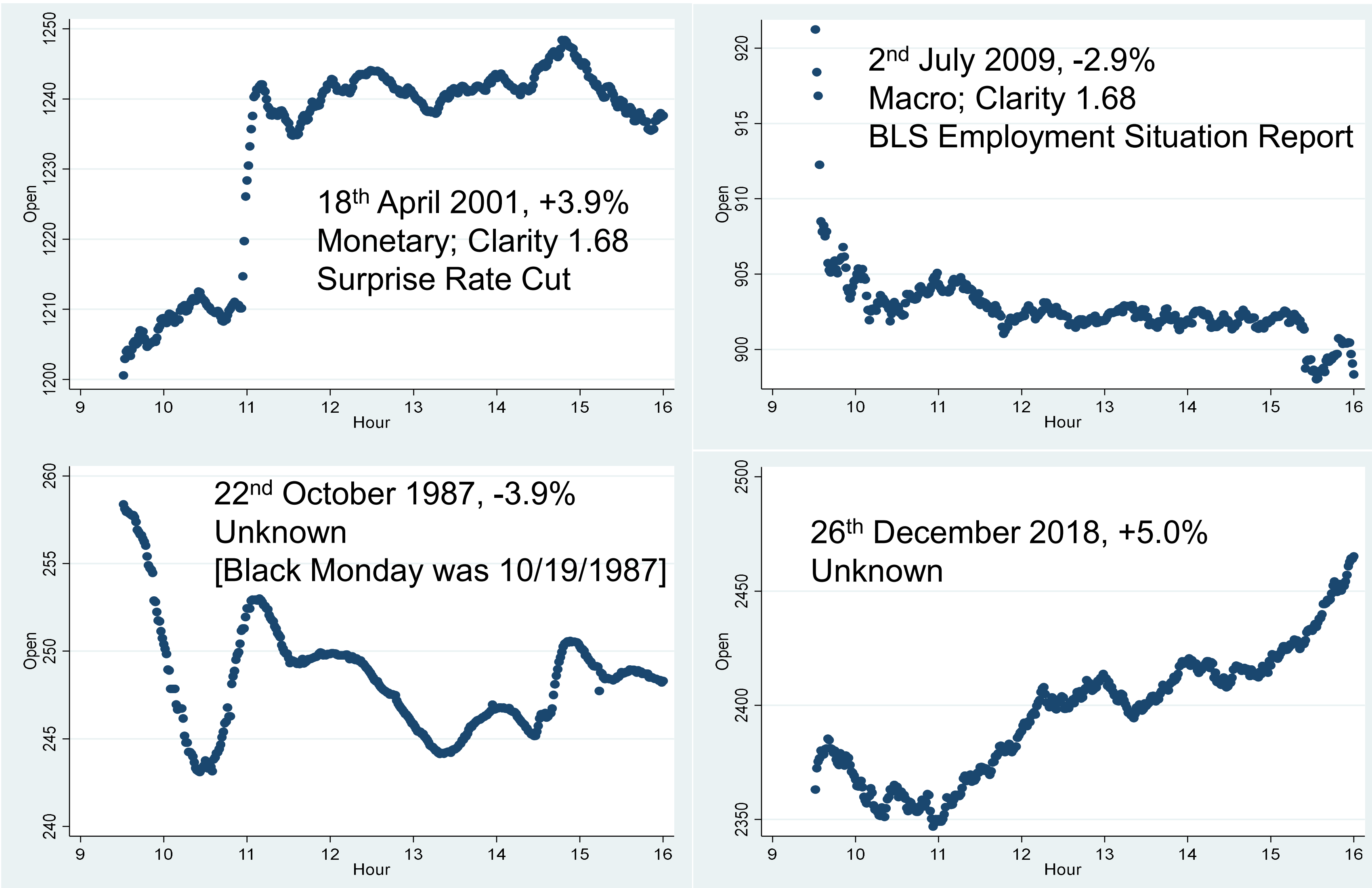
After cleaning the articles, we extract the first 200 words of each article. This has two main benefits: (1) It makes all the articles the same length, which is useful when doing tf-idf to prevent biases caused by differences in document length and (2) many articles, especially early in the sample, discuss 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 model's out of sample performance. 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 slight. 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 S&P Returns and Attributed Driver



Notes: Each panel plots the level of the S&P 500 based on 1-minute increments from market open to market close for the noted dates.

Figure 2: Two Example High Clarity Newspaper Articles

Stock Prices Soar, as Investors Embrace a Surprise Rate Cut

By E.S. BrowningStaff 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.

This WSJ article to the left would receive a primary category of **Monetary Policy (Policy)** because the article links the rise to the surprise interest rate cut. Geographic source would be the **US**. Journalist confidence would be **High**, as the article explicitly links the move to the rate cut. Ease of coding would be **Easy**.

This WSJ article to the right is coded as **Macro News and Outlook (Non-Policy)** because the drop is clearly linked to the poor jobs report. Geographic source would be the **US**. Confidence and ease of coding **High** and **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 Lower Clarity News Paper Articles (from the same day)

This WSJ article below would be coded **Unknown** as it explicitly states there is no explanation for the event.

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.

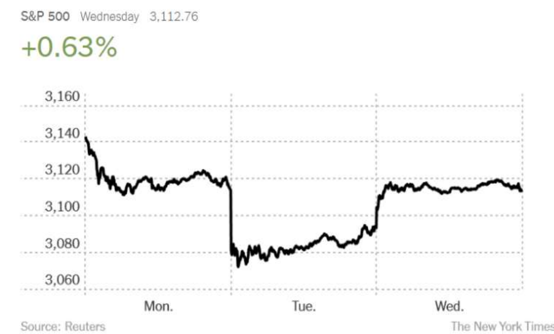
All 30 stocks in the Dow industrials notched gains, as did each of the 11 sectors in the broader S&P. Shares of Amazon.com , Facebook and Netflix climbed more than 8%, while retailers including Kohl's and Macy's rallied as early data on the crucial holiday shopping season appeared robust. Energy stocks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

This article to the right was coded with a primary cause as **Macroeconomic News and Outlook** because the first reason listed was consumer spending. Secondary causes would include **Commodities** and **International Trade Policy**. The geographic source would be the **US**. Confidence and ease of coding are **Moderate** and **Moderate.**

The New York Times

Stocks Bounce Back From Edge of Bear Market



By Emily Flitter
Dec. 26, 2018

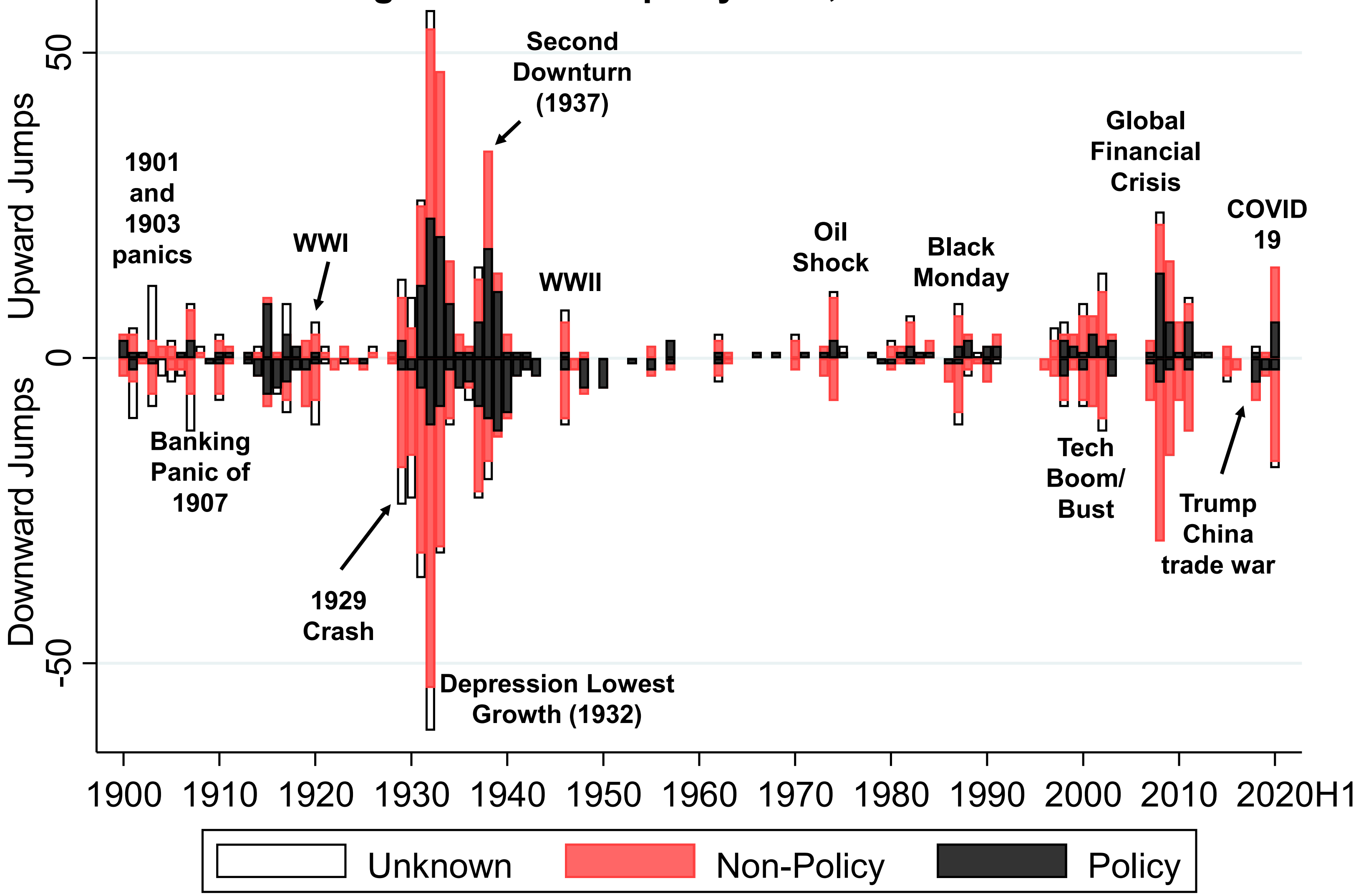
352

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.

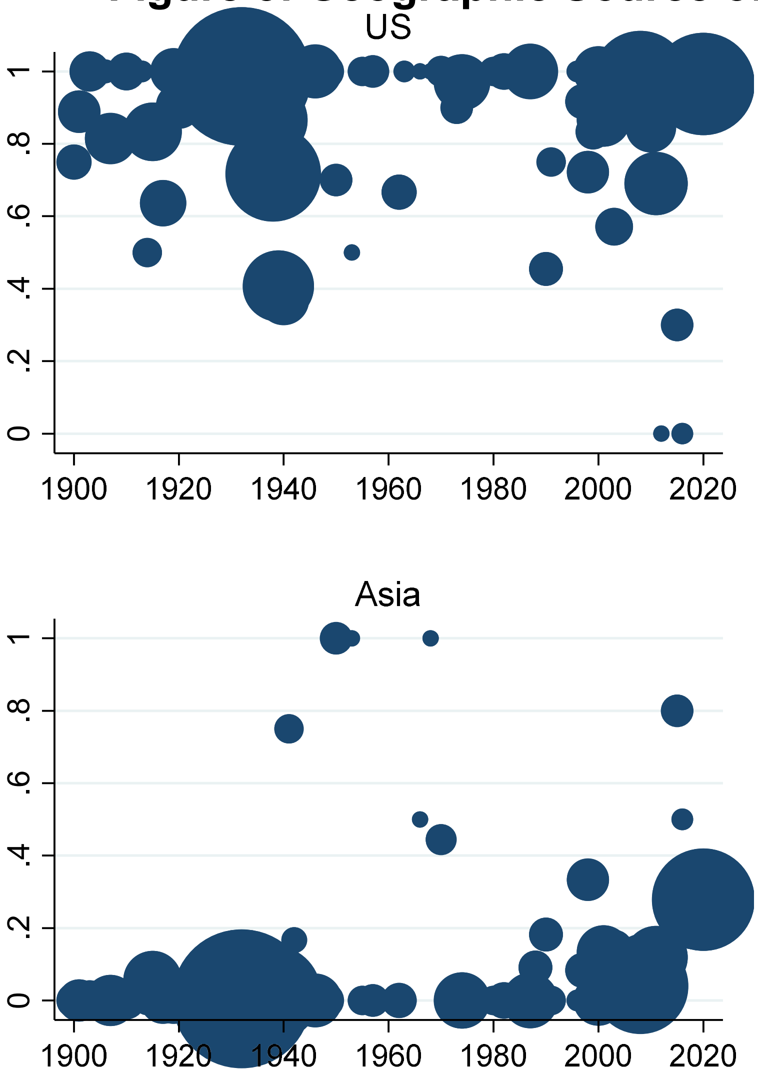
A substantial rise in crude oil prices added to the lighter mood, as did efforts from the White House to ease up on criticism of the Federal Reserve.

Figure 4: US Jumps by Year, 1900-2020



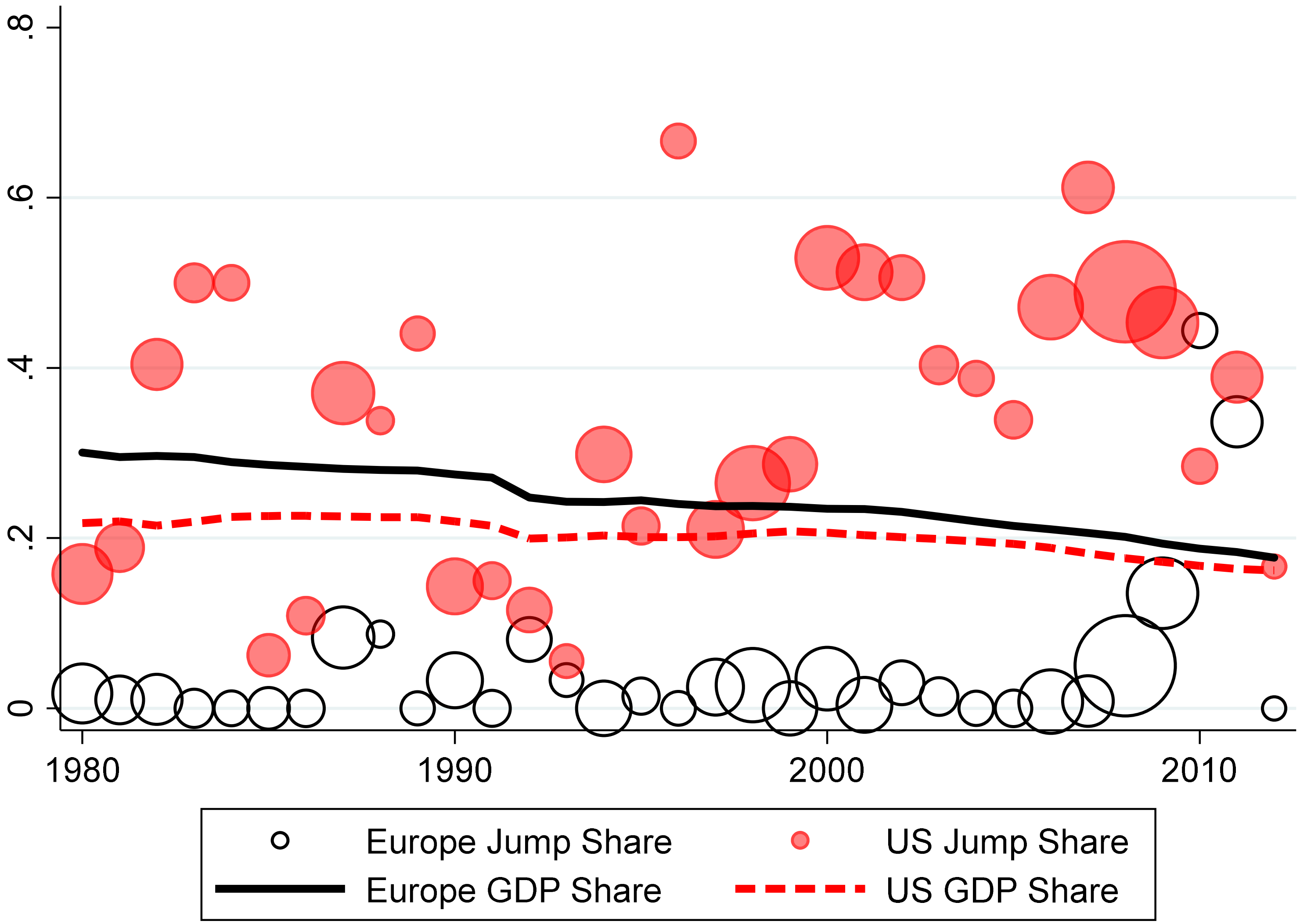
Notes: Each bar is the number of positive or negative jumps in that year. Shadings indicate the number of jumps triggered by "Policy", "Non-Policy" and "Unknown" news. Unknown includes 5 instances of "no article found" between 1900 and 1925. Data from 1900-2020H1.

Figure 5: Geographic Source of US Jumps by Year, 1900-2020



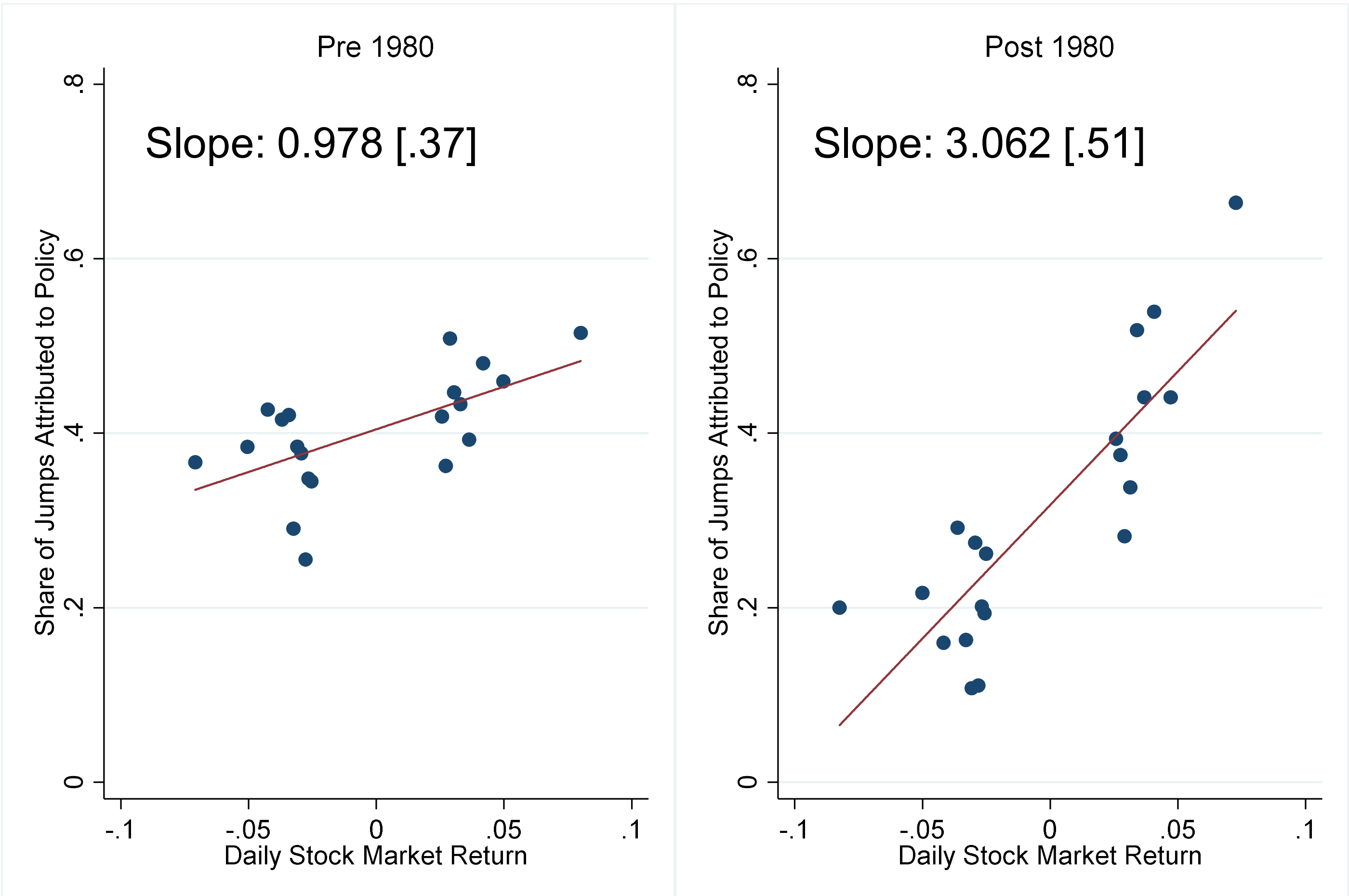
Notes: Dot shows the share of jumps in that year in the US by their geographic origin (US origin jumps top left, European origin jumps top right etc). The size of the dots reflects the number of jumps in that year. Data from 1900 to 2020H1. Excludes unknown and no article found jumps, which have no geographic attribution.

Figure 6: US- and Europe-sourced Share of Jumps in Other Countries, 1980-2012



Notes: Share of US and Europe sourced stock-market jumps averaged over third-party countries by year: Australia, Canada, China (HK), China (Shanghai), Japan, New Zealand, Saudi Arabia, Singapore, South Africa, and South Korea. Dot size proportional to the number of jumps by year. GDP share is the PPP share of world total GDP from the IMF. Data 1980-2012.

Figure 7: Policy-Share by Jump Size and Period, US



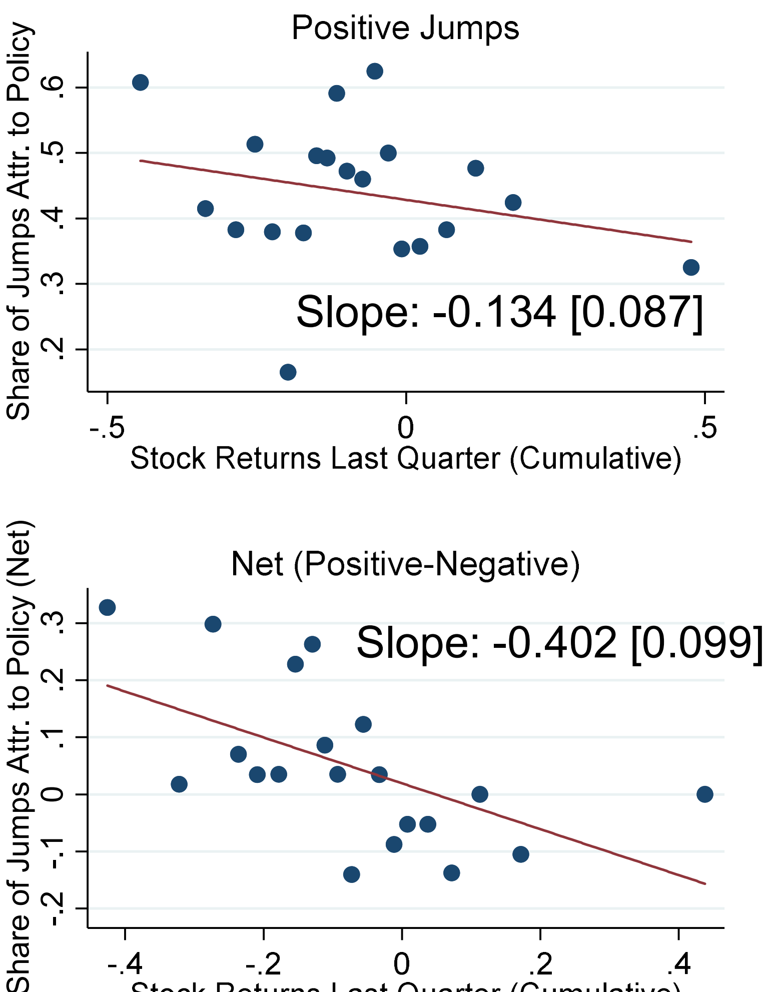
Realized Shares -- Diff. in slopes: 2.08, t-Stat: 9.80.
Fixing Shares -- Diff. in slopes: 1.93, t-Stat: 1.82.

Notes: Plot is a binscatter (n=20) of our policy score on stock returns. For each sub-period, we run a regression of policy on returns, and report the t-Statistic on the return variable. US data 1900-2020H1.
We also regress (for only jump days):

$$policy_t = a + b\ return_t + c\ 1_{post80} + d\ return_t \times 1_{post80} + e_t$$

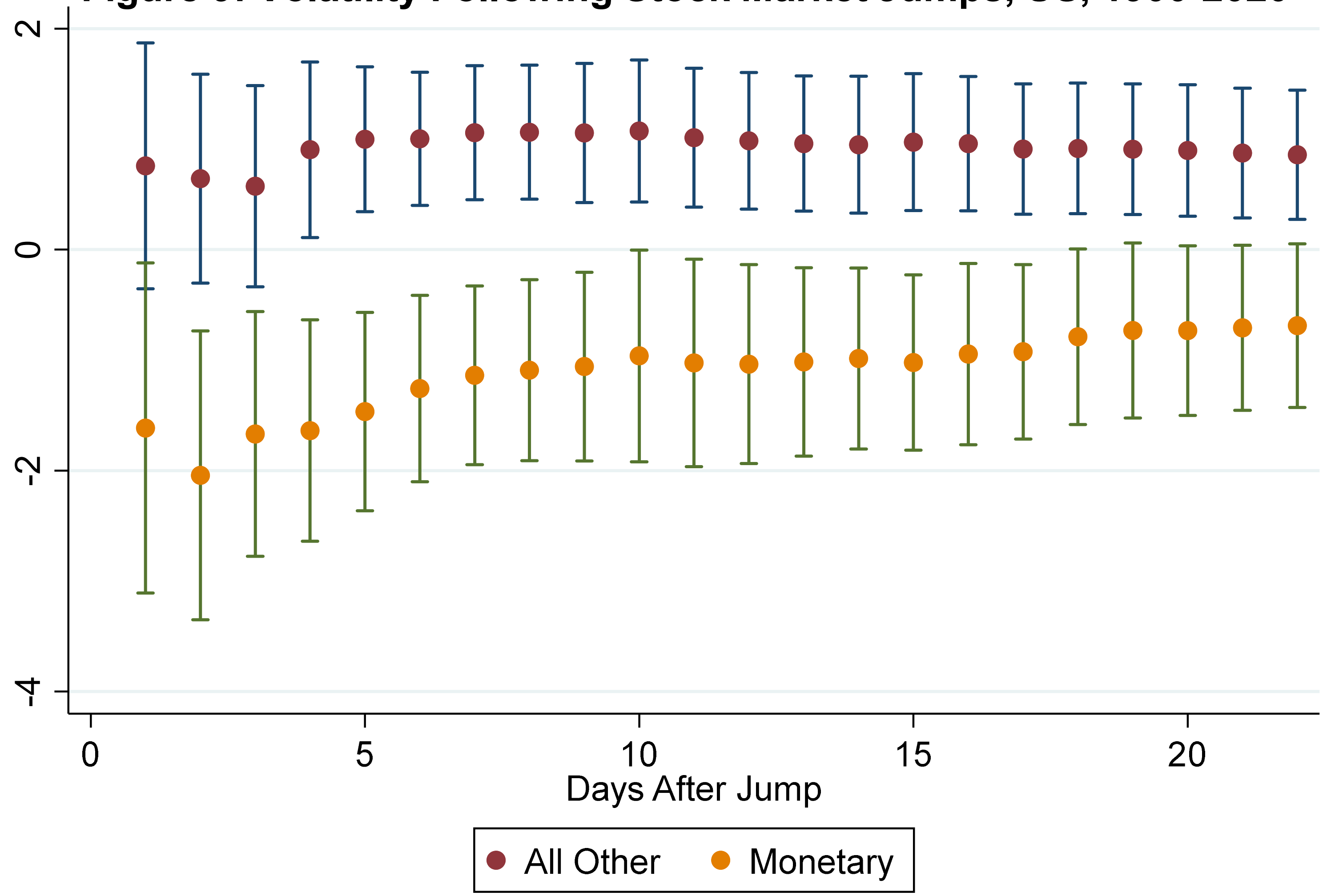
And report the coefficient on the interaction term *d*, and its t-statistic at the bottom of the figure. Of the increase in slope of 2.08 we find that 7.2% (2.08-1.93)/2.08 is due to a shift to categories with a more positive policy mix and the remaining 92.8% is due to policy jumps becoming more positive within each category. Looking at the individual categories we see that they all have a steeper slope after 1980.

Figure 8: Policy and Past Returns



Notes: X-axis is cumulative log returns over the past 66 trading days. For the top two panels, the Y-axis is the score for our policy variable. For the bottom left panel, we compute a new binary policy variable. This takes the value of 1 if the policy score is greater than or equal to 0.5, and zero otherwise. We then create a Net jump policy variable equal to the binary policy variable if the jump is positive, and equal to minus one times the binary policy variable if the jump is negative. Data from the US 1900 to 2020H1. Difference in slopes of 0.575 is statistically significant at the 1% level.

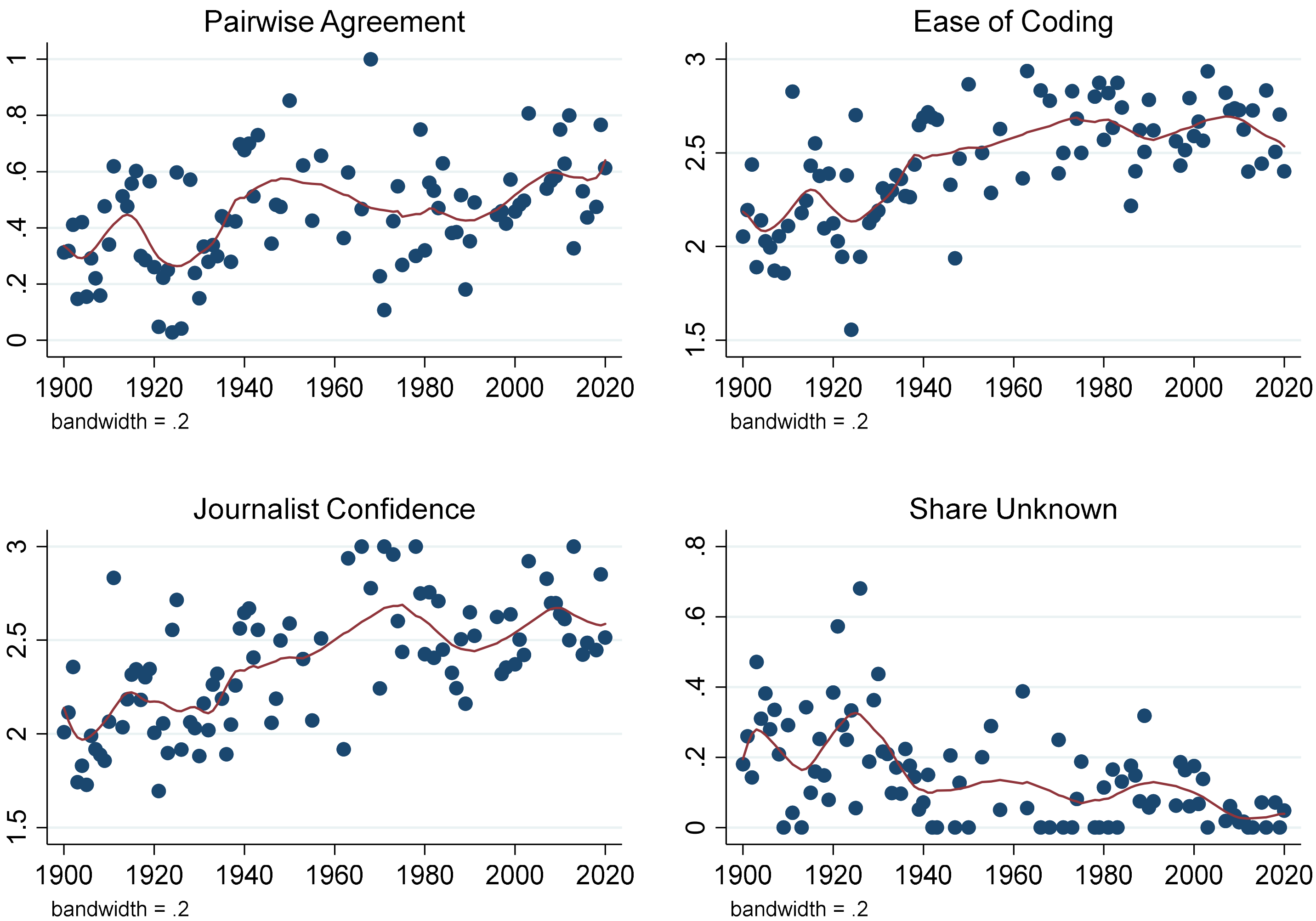
Figure 9: Volatility Following Stock Market Jumps, US, 1900-2020



Bars represent a 95% confidence interval around the point estimate

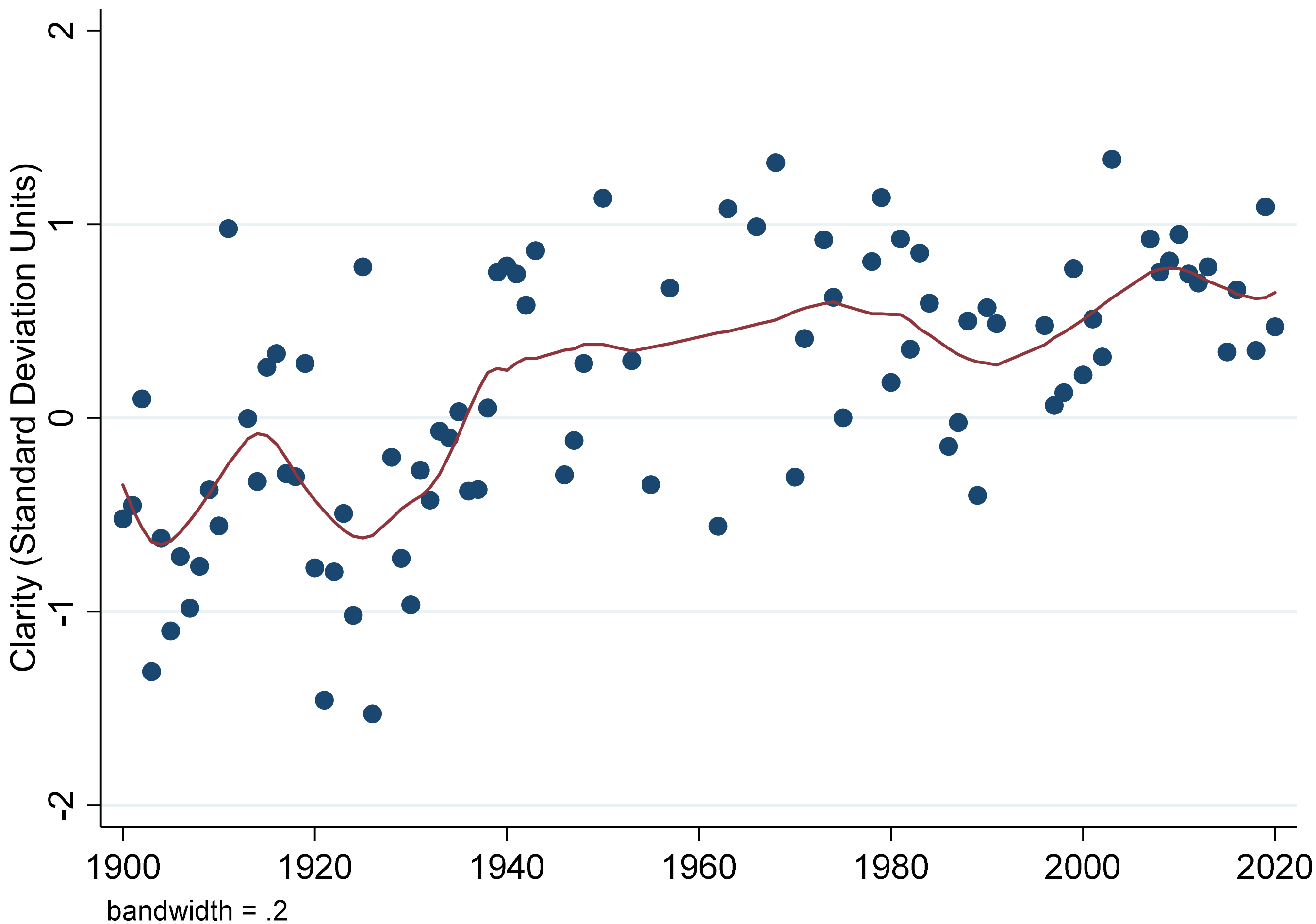
Notes: Volatility is defined as the average squared percentage return over the next *n* days. Controls are the daily return, split into positive and negative components, volatility over the last day, last week and last month (HAR controls). US data, 1900-2020H1. Newey-West standard errors with lags equal to 1.5x the number of overlapping observations.

Figure 10: Clarity Index Components Over Time, US, 1900-2020



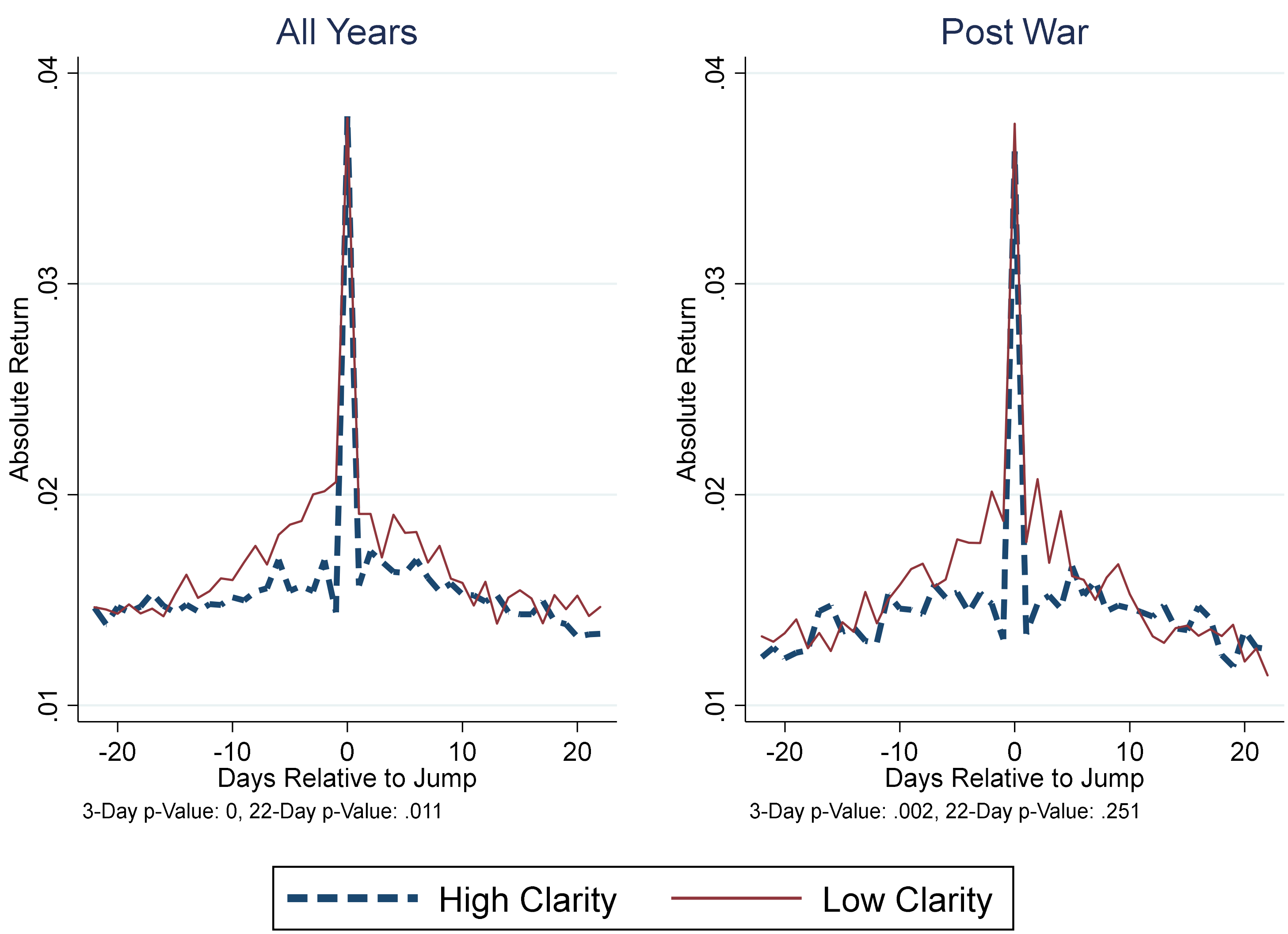
Notes: Red line represents LOWESS smoothing on components of clarity. Bandwidth is share of data used at any point in time to fit the LOWESS polynomial. Pairwise agreement is the average share of pairs of coders that agree (out of up to 45 possible pairs arising from 5 newspapers per day, and two coders per newspaper). Ease of coding is rated on a 1-3 scale, with one being the hardest, and three being the easiest. Journalist confidence is rated on a 1-3 scale, with one being the least confident and three being the most confident. Share unknown is the percentage of coders who marked coded an article as unknown on a given day.
US data, 1900-2020H1.

Figure 11: Clarity Index Over Time, US, 1900-2020



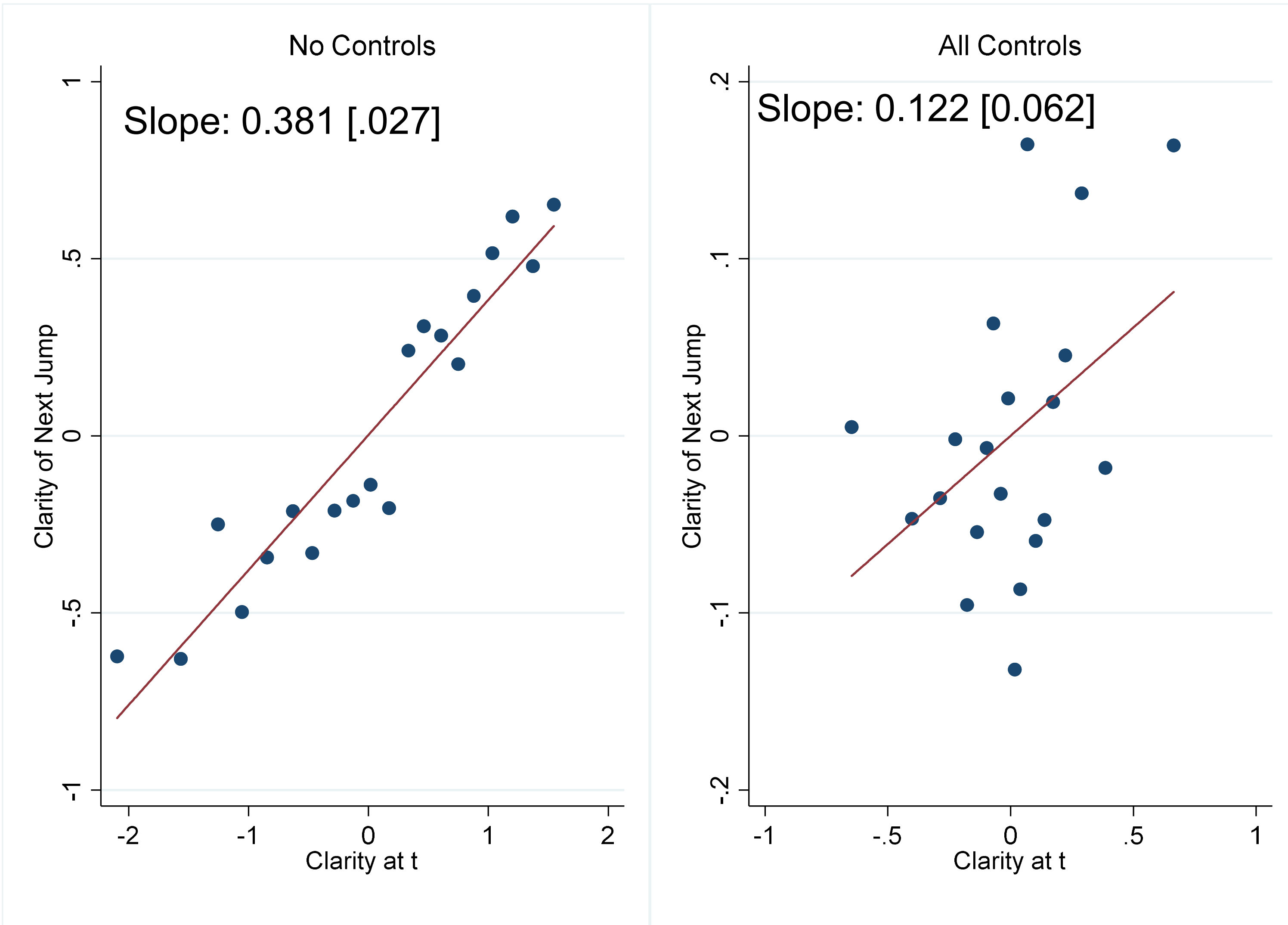
Notes: Red line represents LOWESS smoothing on clarity index. Clarity is the first principal component of: ease of coding, confidence, share of coders who agree and share of "Unknown" codings. It is mean zero and standard deviation one. Bandwidth is share of data used at any point in time to fit the LOWESS polynomial. Pairwise agreement is the average share of pairs of coders that agree (out of up to 45 possible pairs arising from 5 newspapers per day, and two coders per newspaper). Ease of coding is rated on a 1-3 scale, with one being the hardest, and three being the easiest. Journalist confidence is rated on a 1-3 scale, with one being the least confident and three being the most confident. Share unknown is the percentage of coders who marked coded an article as unknown on a given day. US data, 1900-2020H1.

Figure 12: Volatility Around Low and High Clarity Jumps, US, 1900-2020



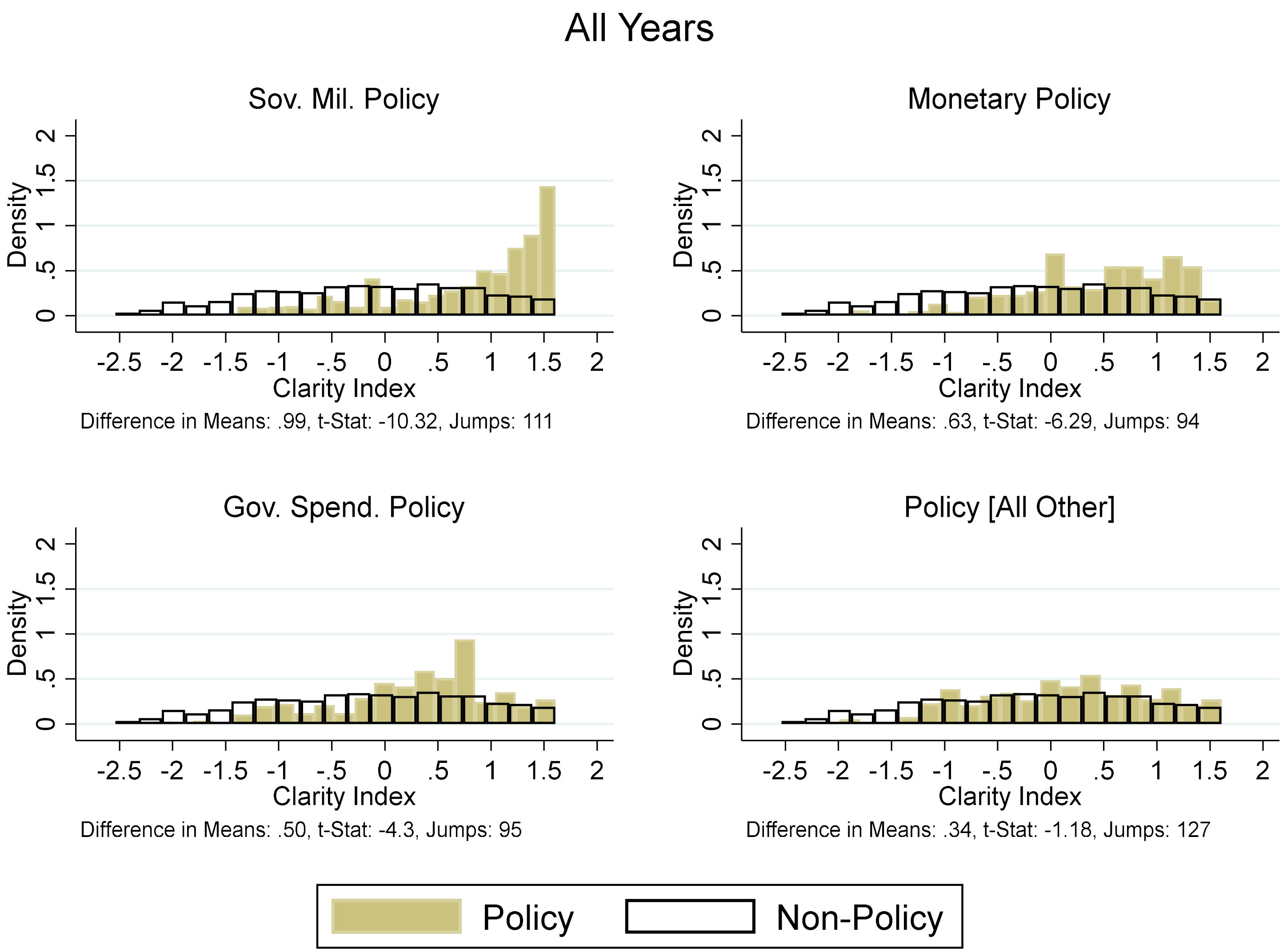
Notes: High (low) clarity is defined as clarity above (below) the sample median for either All Years (1900-2020H1) or Post War (1946-2020H1). Plot is the average absolute return in a +/- 22-day window around each jump. p-Value is from a t-test of whether the mean absolute return in a +/- n-day window around the jump is different between high and low clarity events.

Figure 13: Persistence of Clarity, 1900-2020



Notes: Left panel is a binscatter (n=20) of clarity of a jump, relative to the clarity of the last jump (regardless of how much time has elapsed since that jump). In the right panel, clarity is residualized on the return on the day of the jump (split into positive and negative components), and the HAR controls (volatility over the past day, week and month) before creating the binscatter. US data 1900-2020H1.

Figure 14: Clarity of Selected Policy Categories vs. Non-Policy Categories, 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, 1900-2020H1. Average clarity is higher in every policy subcategory than the average for all non policy subcategories.

Table 1: Jumps by Era and Category, US

	US Equities		UK Equities	ROTW Equities	US Bonds
Time Period:	1900-2020	1946-2020	1930-2020	1980-2012	1970-2013
Macroeconomic News & Outlook	23.8%	32.0%	25.5%	25.4%	56.0%
Corporate Earnings & Outlook	10.8%	12.1%	12.9%	9.4%	1.6%
Sovereign Military & Security Actions	9.7%	5.5%	5.3%	2.6%	0.8%
Monetary Policy & Central Banking	8.2%	10.9%	11.2%	7.8%	24.1%
Government Spending	6.4%	6.4%	8.0%	3.8%	3.5%
Commodities	5.9%	2.9%	2.2%	4.7%	0.5%
Regulation	4.4%	2.4%	5.5%	1.5%	0.0%
Other Non-Policy	4.6%	6.1%	3.9%	10.7%	0.5%
Elections & Political Transitions	2.5%	3.5%	2.6%	1.8%	0.8%
Other Policy	2.3%	1.6%	2.2%	4.0%	0.8%
Taxes	1.9%	1.6%	1.3%	0.3%	1.9%
Exchange Rate Policy & Capital Controls	1.0%	0.7%	1.4%	0.8%	0.3%
International Trade Policy	0.9%	1.2%	0.4%		
Foreign Stock Markets	0.9%	0.6%	4.8%	12.2%	0.0%
Terrorist Attacks & Non-State Violence	0.6%	1.1%	0.5%	1.2%	0.0%
Unknown & No Explanation	15.6%	11.6%	10.3%	11.3%	5.9%
No Article Found	0.4%	0.0%	2.0%	2.4%	3.2%
Total	1,149	437	650	3,593	373

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, and South Korea. ROTW panel is not balanced between 1980 and 2012 (see Appendix Table 1). Data for US/UK stock jumps ends 2020H1. US bond jumps are defined as daily changes in the 10-year treasury yield of more than 15 basis points.

Table 2: Categorical Coding Agreement Rates, US

Time Period	**1900-1945**		**1946-2020**	
	Policy vs. Non-Policy	Granular Categories	Policy vs. Non-Policy	Granular Categories
All Coders & All Papers	75%	41%	80%	53%
All Coders Within Paper	88%	67%	90%	73%
Within WSJ	92%	77%	93%	80%
With Random Assignment	53%	12%	57%	18%

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 H1. 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 94-2020	Elections 1900-2020	Macro 1957-2020
FOMC meeting at t or t-1	2.647***	0.056	-0.049
	(0.328)	(0.058)	(0.258)
Election at t or t-1	-1.262	6.451***	1.581
	(1.637)	(0.266)	(1.447)
Macro Announcement at t	0.287	-0.116	1.062***
	(0.366)	(0.077)	(0.318)
Observations	6,671	32,659	15,985
R-Squared	0.012	0.018	0.001
# Codings in Category	59	37	264
Day of Week FE	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 257 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 2020H1. 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 and the Employment Situation Report (1957-2020), for a total of 1524 known dates. Years vary by column. US data, 1900-2020H1. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. US data, date range varies by column.

Table 4: Clarity and Market Concentration

	% of Total Distance Traveled					
	(1)	(2)	(3)	(4)	(5)	(6)
Clarity	2.437***	2.30***				
	(0.486)	(0.502)				
Avg. Ease of Coding			1.12**			
			(0.557)			
Avg. Confidence				2.14***		
				(0.507)		
Share Unknown					-2.71***	
					(0.568)	
Pairwise Agreement						1.77***
						(0.431)
Observations	320	320	320	320	320	320
R-squared	0.06	0.102	0.064	0.095	0.1	0.098
Return Controls	NO	YES	YES	YES	YES	YES
HAR Controls	NO	YES	YES	YES	YES	YES

Notes: Clarity is the first principal component of: ease of coding, confidence, share of coders who agree and share of "Unknown" codings (multiplied by negative one). It is mean zero and standard deviation one. The left hand side variable is share of total distance traveled in the 5 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, 1985-2020H1. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Clarity and Future Stock Volatility, US

	(1)	(2)	(3)	(4)	(5)	(6)
Clarity	-5.055***	-2.25*				
	(1.57)	(1.24)				
Avg. Ease of Coding			-2.82***			
			(1.07)			
Avg. Confidence				-1.3		
				(1.18)		
Share Unknown					2.58	
					(1.58)	
Pairwise Agreement						-1.05
						(1.32)
Observations	1,146	1,146	1,146	1,146	1,146	1,146
R-squared	0.013	0.246	0.245	0.243	0.247	0.244
Return Controls	NO	YES	YES	YES	YES	YES
HAR Controls	NO	YES	YES	YES	YES	YES

Notes: Left-hand-side is the sum of squared percentage returns over the 5 days following a jump day. Clarity is the first principal component of: ease of coding, confidence, share of coders who agree and share of "Unknown" codings (multiplied by negative one). It is mean zero and standard deviation one. US data, 1900-2020H1. 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	2.50%
Australia	1986-2011	Australian Financial Times	2.50%
Canada	1980-2012	The Globe and Mail	2.00%
China (Hong Kong)	1980-2011	South China Morning Post	3.80%
China (Shanghai)	1994-2013	Shanghai Securities Journal	4.00%
Germany	1987-2012	Handelsblat, FAZ	2.50%
Greece	1989-2015	Kathimerini, To Vima	4.00%
Ireland	1987-2012	The Irish Times	2.50%
Japan	1981-2013	Yomiuri and Asahi	3.00%
New Zealand	1996-2011	New Zealand Herald	2.50%
Saudi Arabia	1994-2013	Al Riyadh	2.50%
Singapore	1980-2013	Business Times and Straits Times	2.50%
South Africa	1986-2013	Business Day	2.50%
South Korea	1981-2011	Chosun Ilbo	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. Jump threshold was chosen such that jumps were approximately 1% of trading days

Table A2. Most Common Reasons for Big Jumps By Era in the United States

	Years	# jumps	Negative Jumps		Positive Jumps	
			Most Common	2nd Most	Most Common	2nd Most
Pre-Fed Era	1900-13	100	Unknown	Corp. Earnings	Unknown	Corp. Earnings
World War I	1914-19	63	Sov. Military	Macro. News	Sov. Military	Unknown
1920s	1920-28	32	Unknown	Macro. News	Unknown	Corp. Earnings
Depression Era	1929-38	466	Macro. News	Unknown	Macro. News	Monetary Policy
World War II	1939-45	51	Sov. Military	Macro. News	Sov. Military	Monetary Policy
Early Postwar	1946-72	63	Macro. News	Sov. Military	Unknown	Sov. Military
Inflation & Oil shocks	1973-79	27	Macro. News	Commodities	Monetary Policy	Macro. News
Disinflation & Growth	1980-94	65	Macro. News	Corp. Earnings	Macro. News	Monetary Policy
Boom, Rec. & Recovery	1995-2006	95	Macro. News	Corp. Earnings	Monetary Policy	Macro. News
Global Financial Crisis	2007-10	109	Macro. News	Corp. Earnings	Macro. News	Corp. Earnings
Post GFC	2011-18	37	Macro. News	Unknown	Macro. News	Monetary Policy
All Periods	**1900-2018**	**1108**	**Macro. News**	**Unknown**	**Macro. News**	**Unknown**

Notes: We identify the 10 biggest stock market gains/losses in each era, and identify the modal categories among these moves.

Table A3: Breakdown of Policy Jumps by Country

	# Policy Jumps	
	Positive	Negative
Australia	15	12
Canada	52	49
Germany	46	39
Greece	50	33
Hong Kong	42	23
Ireland	57	43
Japan	62	39
Korea	67	42
New Zealand	2	1
Saudi Arabia	25	21
Shanghai	52	51
Singapore	49	40
South Africa	46	44
UK	26	20
All	591	457

Notes: *Positive* (*Negative*) columns are share of positive (negative) jumps attributed to policy categories. Unbalance panel, see Table A.1 for sample period for each country.

Table A4: Regression Models Fit to Daily Industry-Level Equity Returns from 1960 to 2016

	Banks		*Pooled Sample*	
	(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 A5: Policy-Share by Jump Size and Period, US

Absolute Jump Size	US				Rest of the World	
	1900-1979		1980-2020		1980-2015	
	Positive	Negative	Positive	Negative	Positive	Negative
[2.5%,3%)	43%	33%	36%	20%	24%	23%
[3%,3.5%)	42%	35%	43%	12%	38%	16%
[3.5%,4%)	40%	43%	48%	33%	32%	23%
4% or Larger	48%	39%	54%	19%	45%	27%
All	45%	33%	43%	20%	37%	24%
p-Value	0.01		0.00		0.00	
Total	802		347		5,637	

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 1900-2020H1.

Table A6: Policy-Share by Jump Size, Period and Category, US

	1900-1979		1980-2020		Slope	p-Value
	Positive	Negative	Positive	Negative		
Policy	**164**	**156**	**73**	**35**	**0.022**	**0.000**
Sovereign Military & Security Actions	34	66	5	6	0.011	0.410
Monetary Policy & Central Banking	36	17	30	11	0.009	0.296
Government Spending	34	12	18	10	(0.009)	0.488
Regulation	20	28	3	1	0.027	0.239
Taxes	7	10	5	0	0.035	0.015
All Other Policy	33	24	12	8	0.006	0.603
Non Policy	**146**	**193**	**70**	**131**	**(0.029)**	**0.000**
Macroeconomic News & Outlook	73	81	40	79	(0.030)	0.000
Corporate Earnings & Outlook	33	44	21	26	0.002	0.855
Commodities	25	37	2	4	(0.015)	0.502
All Other Non-Policy	16	31	8	22	(0.014)	0.246

Notes: Results shown for most common policy and non-policy categories. The final two columns are the from the coefficient d in the regression: $return_t = a + b\ category_{\mathrm{t}} + c\ 1_{post80} + d\ category_t\ \times 1_{post80} + e_{\mathrm{t}}$ US data, 1900-2020H1.

Table A7: Volatility Following Policy and Non-Policy Jumps, US, 22-day

		Next 22 Days				Next 5 Days	
		(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	Policy	3.534***	0.11				
		(0.539)	(0.321)				
	Non-Policy	4.774***	0.916**				
		(0.667)	(0.365)				
Non-Policy	Commodities			6.667***	1.447*	8.494***	1.755
				(1.210)	(0.877)	(2.167)	(1.382)
	Corporate Earnings			3.374***	0.409	3.581***	-0.438
				(0.778)	(0.507)	(1.299)	(0.837)
	Macro News			5.148***	1.418***	6.265***	1.351
				(0.843)	(0.530)	(1.158)	(0.926)
Policy	Monetary Policy			2.069***	-0.911*	1.825***	-2.028***
				(0.579)	(0.526)	(0.616)	(0.569)
	Fiscal Policy			6.501***	1.595	7.234***	0.956
				(1.547)	(1.251)	(1.816)	(1.171)
	Sovereign Military			1.526***	-0.662*	3.329***	0.176
				(0.413)	(0.402)	(0.976)	(0.882)
	Obs	32,238	32,216	32,238	32,216	32,255	32,233
	R-Squared	0.11	0.321	0.119	0.323	0.108	0.314
	Return Controls	NO	YES	NO	YES	NO	YES
	HAR Controls	NO	YES	NO	YES	NO	YES
F-Test for joint equality of coefs.		0.0267	0.0715	0.00006	0.0115	6.34E-05	0.00473

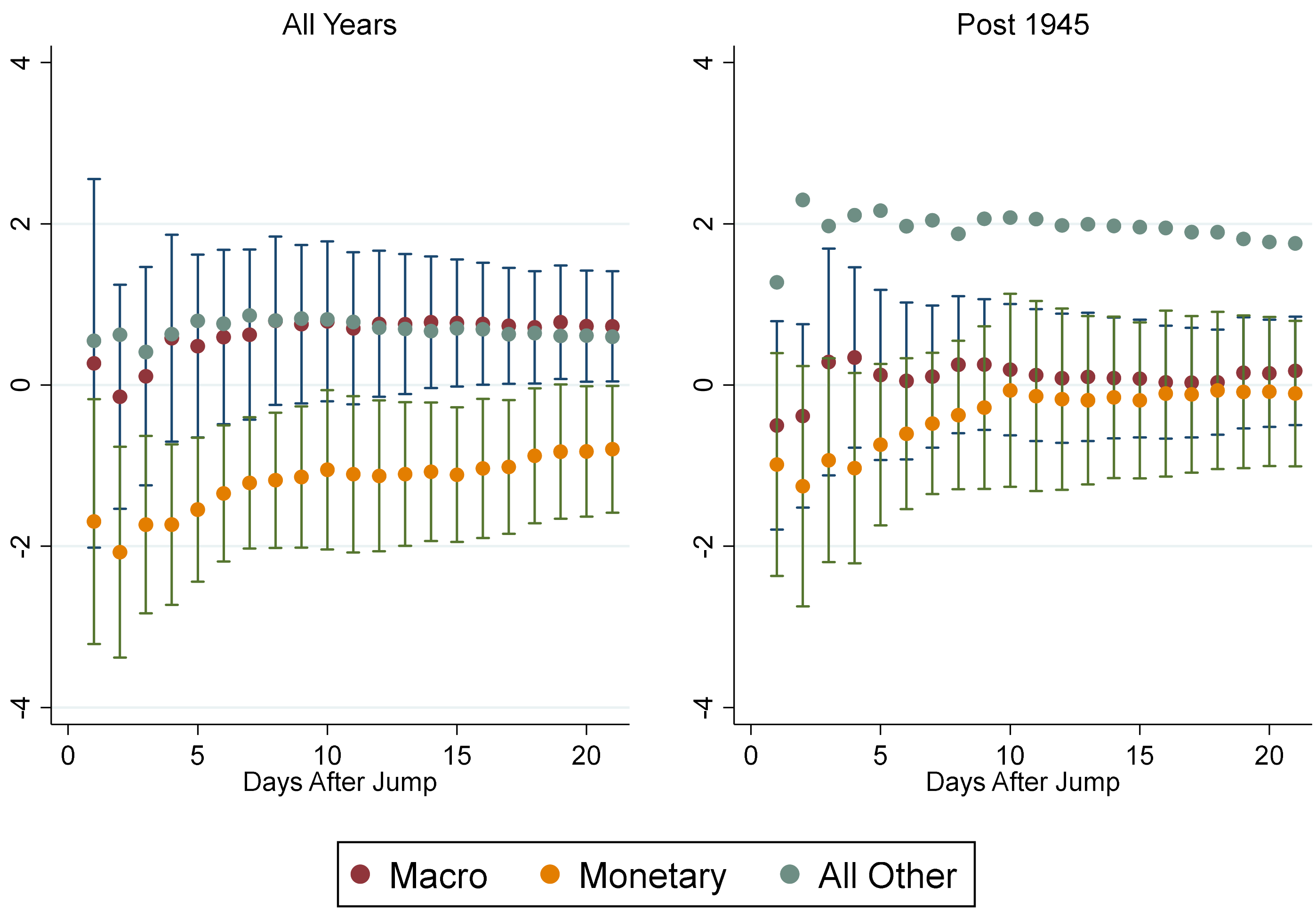
Notes: Columns 1-4 represent regressions, where the left-hand-side is the average percentage squared return in the 22 days following the jump. In columns 5-6, the left-hand-side is the average percentage squared return in the 5 days following the jump. US data, 1900-2020H1. There are only dummy variables 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

Table A8: Clarity Persistence, US, 1900-2020

	LHS: Clarity of Jump at t (All Years)						
Clarity of Last Jump	0.244***	0.195***	0.136***	0.137***	0.111***	0.116***	0.118***
	(0.0305)	(0.0405)	(0.0302)	(0.0298)	(0.0304)	(0.0302)	(0.0306)
Return of Last Jump (Positive)				-4.265***		-3.044**	-3.027**
				(1.1850)		(1.3410)	(1.3410)
Return of Last Jump (Abs, Negative)				-5.680***		-4.744***	-4.816***
				(1.1700)		(1.3060)	(1.3130)
Volatility last day					-8.512	-9.241	-8.991
					(11.8700)	(11.3400)	(11.3300)
Volatility last week					-9.146	-5.552	-5.338
					(5.8450)	(5.9900)	(6.0040)
Volatility last month					-1.272	-1.286	-1.096
					(2.0430)	(2.0400)	(2.0480)
Time Since Last Jump							0.000196
							(0.0001)
Time Since Last Jump x Clarity							-0.0000419
							(0.0001)
Observations	1,146	1,146	1,146	1,146	1,144	1,144	1,144
R-squared	0.23	0.265	0.592	0.598	0.598	0.603	0.603
Specification	Time Trend, Postwar, Interaction	Add Cat. Of Last Jump	Add Cat of Current Jump	Return Controls	Har Controls	Return/Har controls	Time since last jump

Notes: Regressions, where the left-hand-side is the clarity of a jump on date t, 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 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 time since the last jump (in days) and interacts this with the clarity of the last jump.US data, 1900-2020H1. Robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

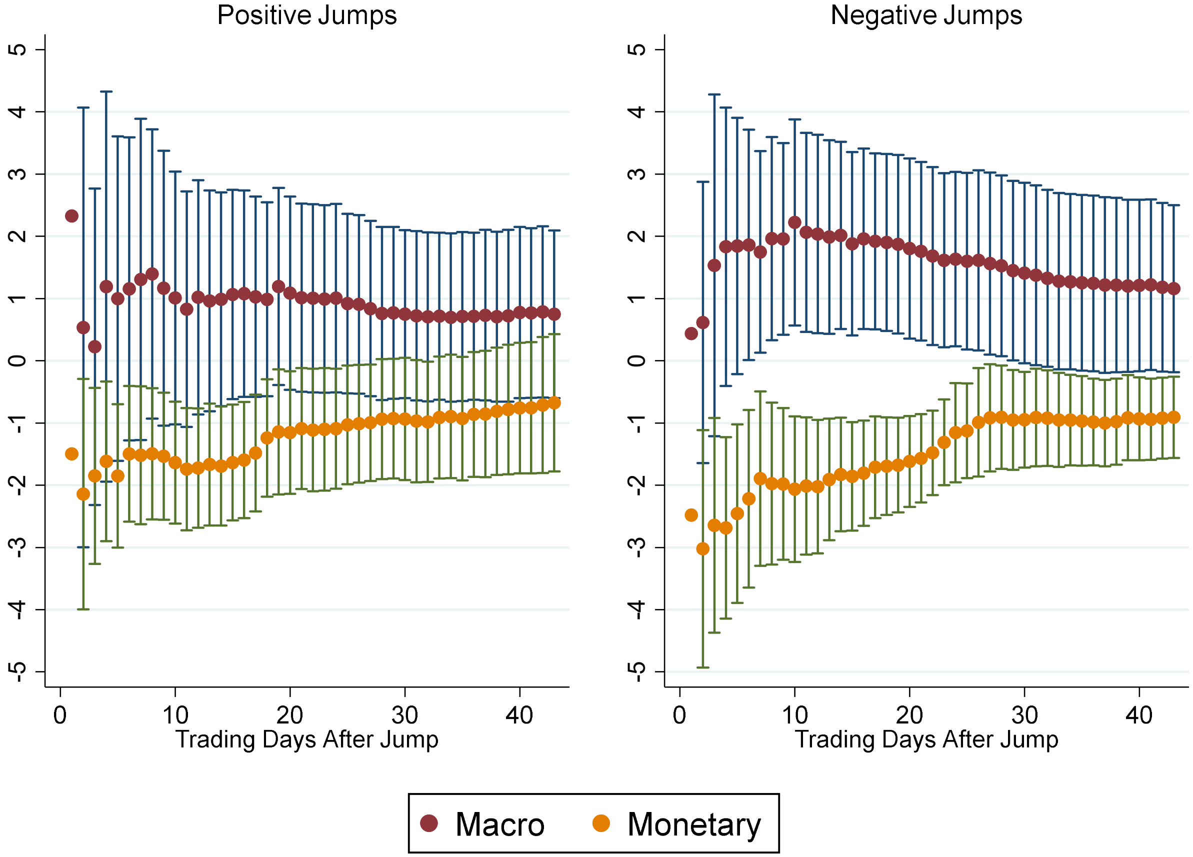
Figure A1: Volatility Following Macro/Monetary



Bars represent a 95% confidence interval around the point estimate

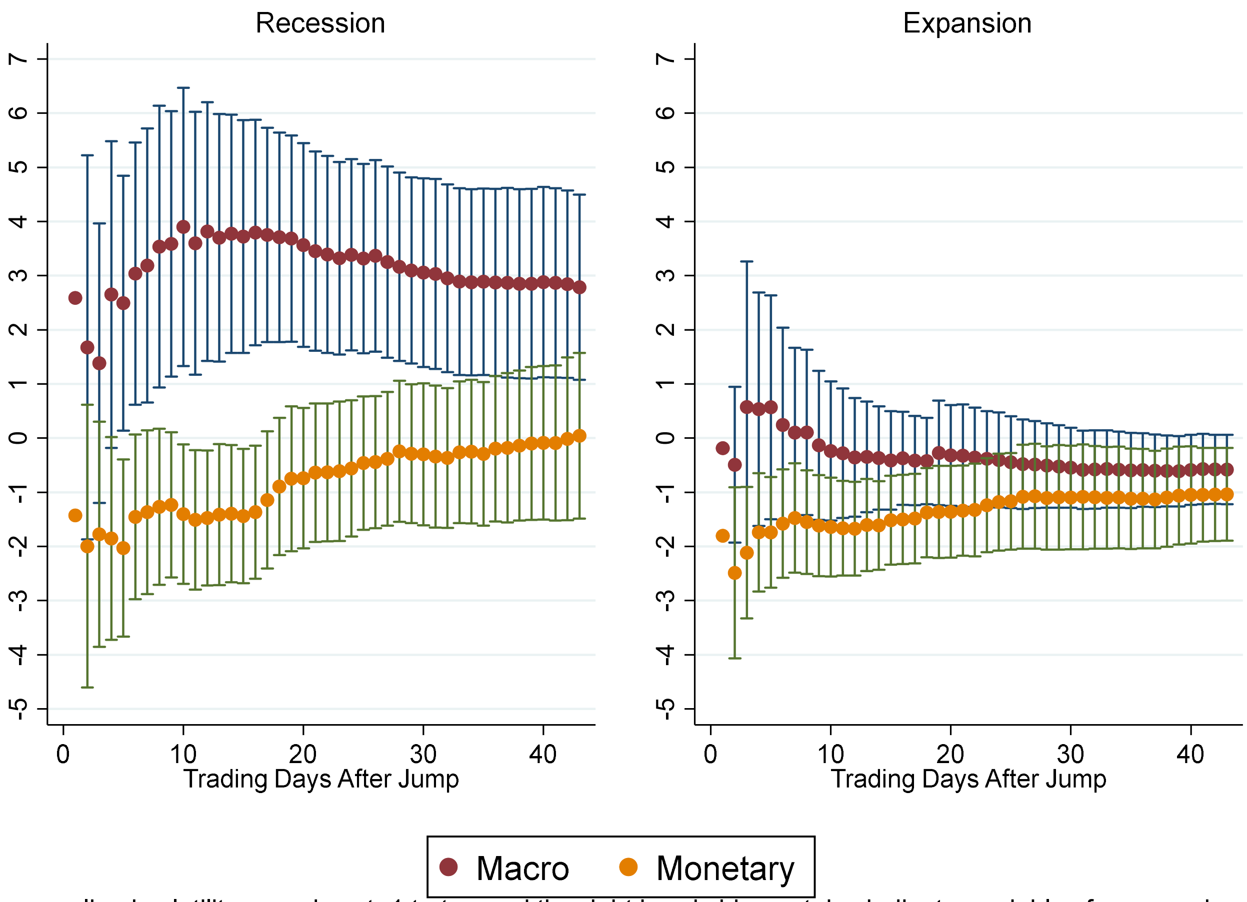
Notes: Volatility is defined as the average squared percentage return over the next *n* days. Controls are the daily return, split into positive and negative components, volatility over the last day, last week and last month (HAR controls). US data, 1900-2020H1. Newey-West standard errors with lags equal to 1.5x the number of overlapping observations.

Figure A2: Volatility Following Jumps, Positive and Negative Jumps, 1900-2019



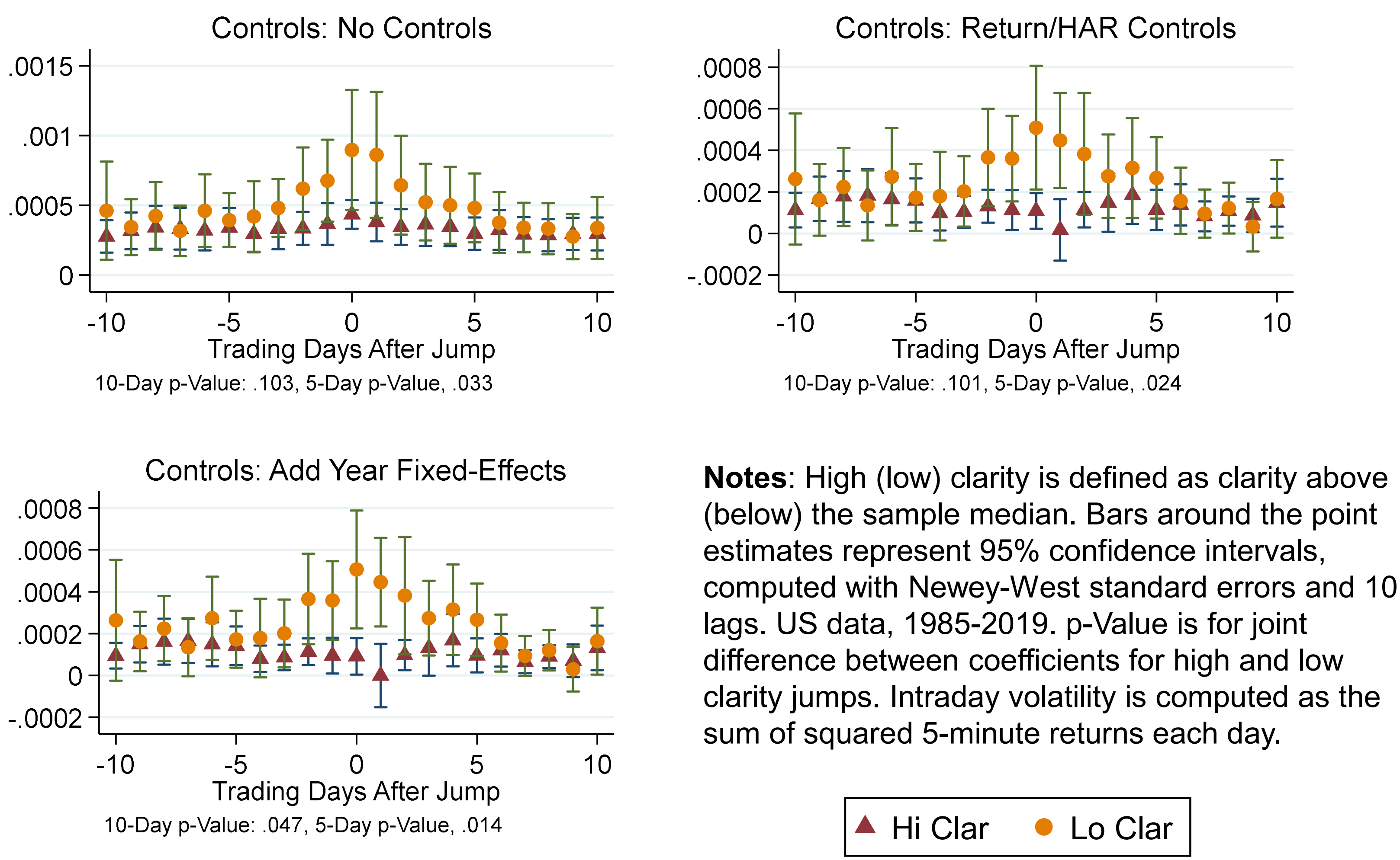
Notes: Left hand side is average realized volatility over days t+1 to t+n. Right hand side is an indicator variable for macro jumps with returns greater than zero, macro jumps with returns less than zero, etc. We also include HAR controls for volatility over the last day, week, and month. Bars represent a 95% confidence around the point estimate computed with Newey-West standard errors and lags equal to 1.5x the number of overlapping observations. US data, 1900-2019. Standard error bars omitted for n=1 to simplify scaling.

Figure A3: Volatility Following Jumps, Recessions and Expansions, 1900-2019



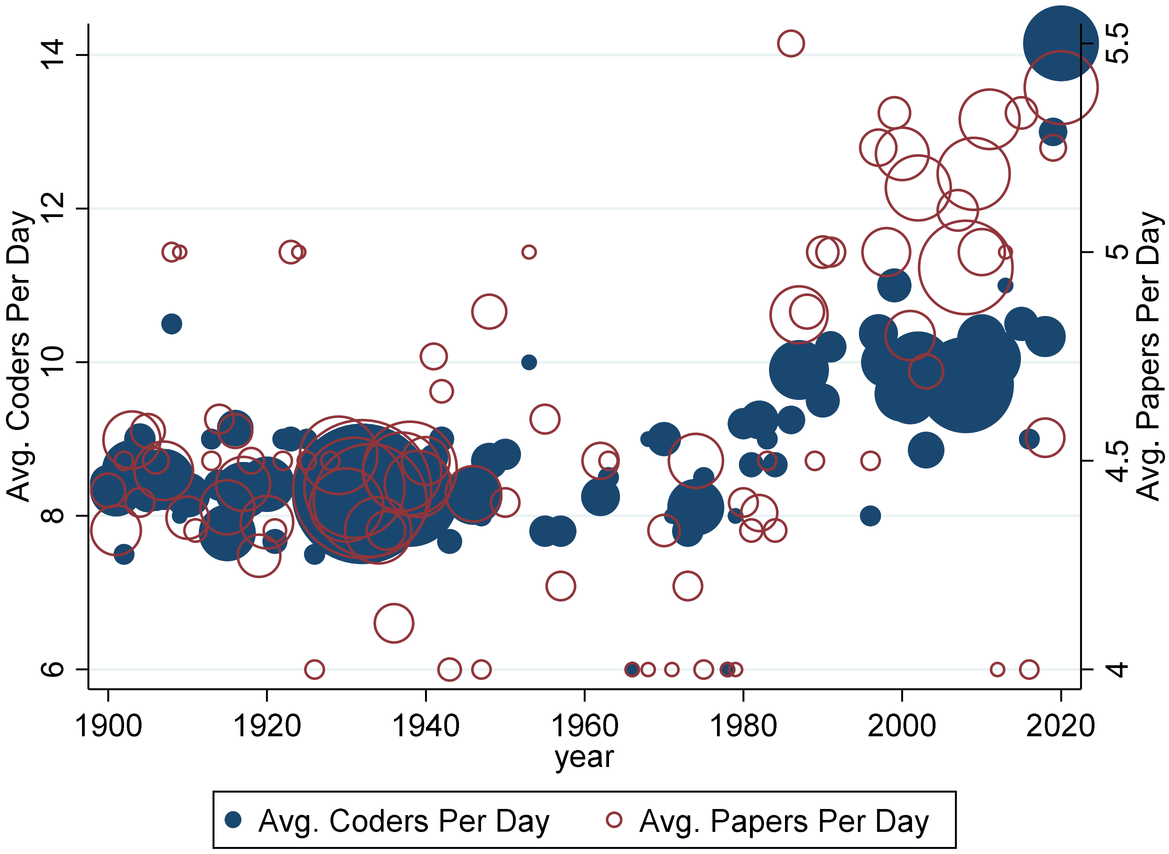
Notes: The left hand side is average realized volatility over days t+1 to t+n and the right hand side contains indicator variables for macro jumps that occur during NBER recessions, macro jumps that occur outside of NBER recessions, etc. We also include HAR controls for volatility over the last day, week, and month. Bars represent a 95% confidence around the point estimate computed with Newey-West standard errors and lags equal to 1.5x the number of overlapping observations. US data, 1900-2019. Standard error bars omitted for n=1 to simplify scaling.

Figure A4: Clarity and Intraday Volatility 10 around 1985-2019



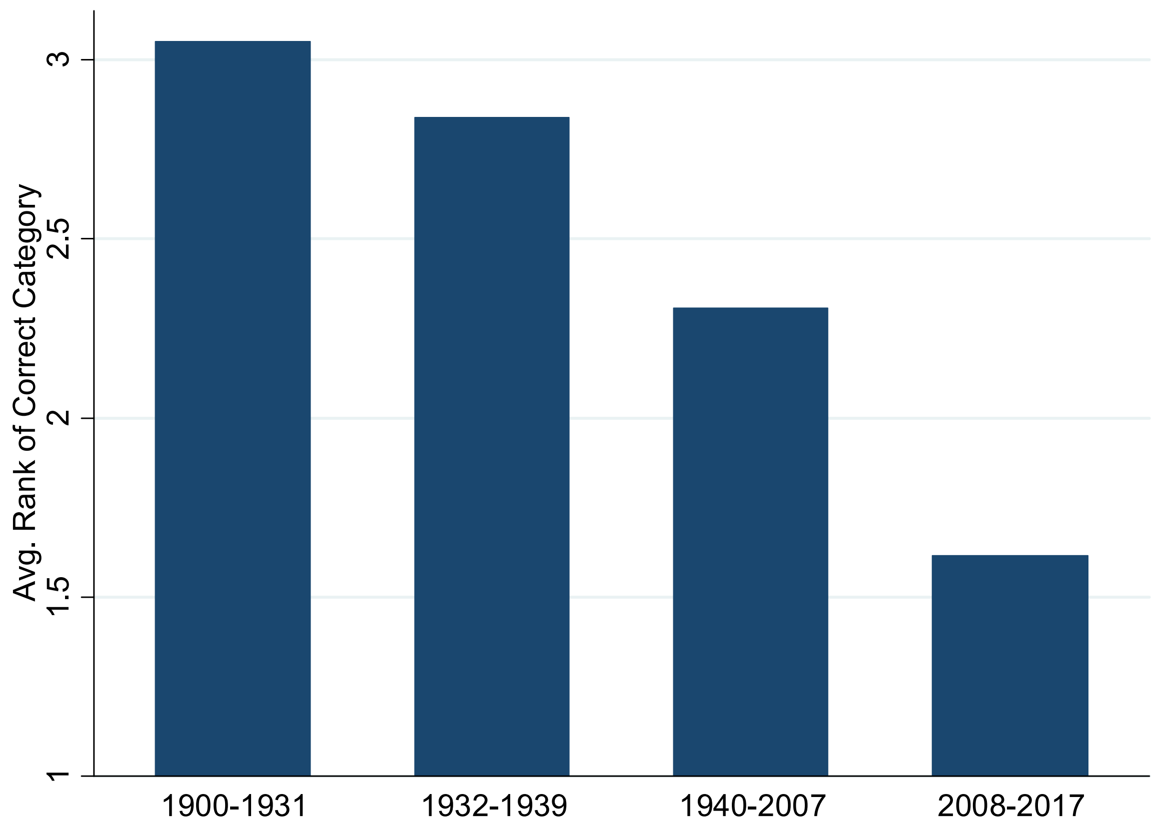
Notes: High (low) clarity is defined as clarity above (below) the sample median. Bars around the point estimates represent 95% confidence intervals, computed with Newey-West standard errors and 10 lags. US data, 1985-2019. p-Value is for joint difference between coefficients for high and low clarity jumps. Intraday volatility is computed as the sum of squared 5-minute returns each day.

Figure A5: Average coders and newspaper per day by year



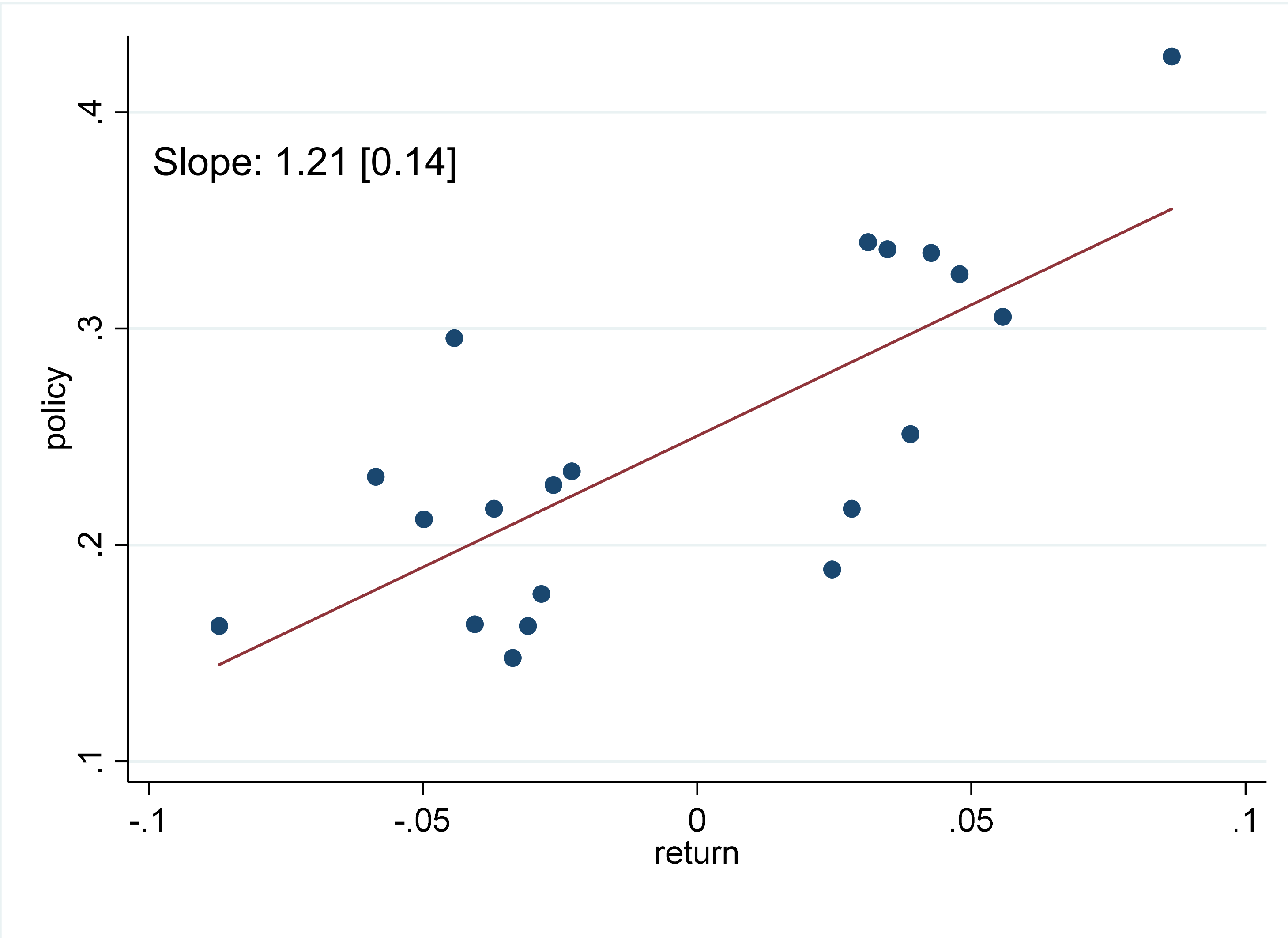
Notes: Shows average number of coders and newspaper per day, with the circle areas proportional to the number of jumps in that year. Data from 1900 to 2020H1

Figure A6: Out-of-Sample Algorithmic Ranking



Notes: 275 jumps in each period. After cleaning/stemming articles 3K unique words remain. Take top 100 words for each category, then add up tf-idf scores for each word for each category in each article. To clean the articles, we take the first 200 words in the article, require words appear in a category at least 3 times, and overall at least 5 times, take top 100 words by tf idf within each article. Exclude ‘Other’ and ‘Unknown’, as well as categories that do not appear at least 5 times in each sub-sample. Out of sample is based on a leave-one-out approach.

Figure A7: Policy-Share by Jump Size and Period, ROTW



Notes: Plot is a binscatter of our policy score on stock returns. For each sub-period, we run a regression of our policy score on returns, and report the t-Statistic on the return variable. Non-US data, 1980-2012.

Table B1: Jumps by Era and Category, UK

<table>
<tr><th></th><th colspan="4">UK Data</th></tr>
<tr><th>Time Period:</th><th>1930-2020</th><th>1930-1945</th><th>1946-2020</th><th>1980-2015</th></tr>
<tr><td>Macroeconomic News & Outlook</td><td>25.5%</td><td>6.3%</td><td>27.0%</td><td>30.9%</td></tr>
<tr><td>Corporate Earnings & Outlook</td><td>12.9%</td><td>1.6%</td><td>13.8%</td><td>13.4%</td></tr>
<tr><td>Sovereign Military & Security Actions</td><td>5.3%</td><td>34.7%</td><td>3.0%</td><td>4.0%</td></tr>
<tr><td>Monetary Policy & Central Banking</td><td>11.2%</td><td>2.2%</td><td>11.9%</td><td>12.6%</td></tr>
<tr><td>Government Spending</td><td>8.0%</td><td>6.9%</td><td>8.0%</td><td>7.2%</td></tr>
<tr><td>Commodities</td><td>2.2%</td><td>1.5%</td><td>2.2%</td><td>2.2%</td></tr>
<tr><td>Regulation</td><td>5.5%</td><td>1.9%</td><td>5.7%</td><td>2.6%</td></tr>
<tr><td>Other Non-Policy</td><td>3.9%</td><td>1.6%</td><td>4.1%</td><td>5.1%</td></tr>
<tr><td>Elections & Political Transitions</td><td>2.6%</td><td>4.6%</td><td>2.5%</td><td>1.7%</td></tr>
<tr><td>Other Policy</td><td>2.2%</td><td>4.7%</td><td>2.0%</td><td>1.5%</td></tr>
<tr><td>Taxes</td><td>1.3%</td><td>2.8%</td><td>1.2%</td><td>0.4%</td></tr>
<tr><td>Exchange Rate Policy & Capital Controls</td><td>1.4%</td><td>1.1%</td><td>1.4%</td><td rowspan="2">1.7%</td></tr>
<tr><td>International Trade Policy</td><td>0.4%</td><td>0.5%</td><td>0.4%</td></tr>
<tr><td>Foreign Stock Markets</td><td>4.8%</td><td>1.4%</td><td>5.1%</td><td>6.7%</td></tr>
<tr><td>Terrorist Attacks & Non-State Violence</td><td>0.5%</td><td>0.0%</td><td>0.5%</td><td>0.9%</td></tr>
<tr><td>Unknown & No Explanation</td><td>10.3%</td><td>27.0%</td><td>9.1%</td><td>6.6%</td></tr>
<tr><td>No Article Found</td><td>2.0%</td><td>1.1%</td><td>2.0%</td><td>2.1%</td></tr>
<tr><td>Total</td><td>650</td><td>46</td><td>604</td><td>340</td></tr>
</table>

Notes: Data for UK runs 1930-2020H1. Jump threshold is 2.5%. From 1930-1983 we use GFD's "UK Industrials" index. From 1984-1993 we use the percent change in the FTSE 100 index level. From 1994-2020 we use the FTSE 100 total return index.

Table B1: Categorical Coding Agreement Rates, UK

Time Period	1930-1945		1946-2020	
	Policy vs. Non-Policy	Granular Categories	Policy vs. Non-Policy	Granular Categories
All Coders & All Papers	68%	46%	74%	45%
All Coders Within Paper	77%	66%	82%	56%
Within the FT	57%	57%	84%	53%
With Random Assignment	52%	21%	54%	13%

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 Financial Times, the Telegraph and the Times of London. UK data runs 1930-2020H1. 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 B3: Policy-Share by Jump Size and Period, UK

Absolute Jump Size	UK 1930-1979 Positive	UK 1930-1979 Negative	UK 1980-2020 Positive	UK 1980-2020 Negative
[2.5%,3%)	36%	43%	36%	26%
[3%,3.5%)	50%	42%	49%	28%
[3.5%,4%)	46%	44%	50%	18%
4% or Larger	58%	38%	48%	11%
All	45%	33%	43%	20%
p-Value	0.31		0.00	
Total	310		343	

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. UK data 1930-2020H1.

Table B4: Policy-Share by Jump Size, Period and Category, UK

	1930-1979		1980-2020		p-Value	Fischer's
	Positive	Negative	Positive	Negative	from t-Test	Exact Test
Policy	**72**	**65**	**67**	**42**	**0.055**	**0.069**
Sovereign Military & Security Actions	6	9	4	6	0.853	1.000
Monetary Policy & Central Banking	8	9	19	6	0.009	0.002
Government Spending	9	7	10	4	0.225	0.761
Regulation	8	7	4	1	0.019	0.371
Taxes	3	1	0	0	0.951	1.000
Non Policy	**53**	**76**	**71**	**137**	**0.121**	**0.211**
Macroeconomic News & Outlook	12	22	11	54	0.006	0.011
Corporate Earnings & Outlook	7	16	9	13	0.263	0.150
Commodities	3	1	2	1	0.557	1.000

Notes: Results shown for most common policy and non-policy categories. The final column with the t-Stat is from a test of whether the share of positive jumps among each category is higher in 1980-2020H1 than 1930-1979.

Table B5: Volatility Following Policy and Non-Policy Jumps, UK, 22-day

		(1)	(2)	(3)	(4)
	Policy	2.927***	0.08		
		(0.778)	(0.419)		
	Non-Policy	3.466***	0.272		
		(0.490)	(0.248)		
Non-Policy	Commodities			3.191*	0.13
				(1.679)	(1.319)
	Corporate Earnings			4.002***	0.987
				(0.959)	(0.705)
	Macro News			3.120***	-0.0247
				(0.604)	(0.292)
Policy	Monetary Policy			3.194**	0.419
				(1.430)	(0.864)
	Fiscal Policy			4.617***	0.926
				(1.367)	(0.799)
	Sovereign Military			1.651***	-0.949*
				(0.484)	(0.504)
	Obs	23,059	23,037	23,059	23,037
	R-Squared	0.094	0.368	0.096	0.369
	Return Controls	NO	YES	NO	YES
	HAR Controls	NO	YES	NO	YES
F-Test for joint equality of coefs.		0.429	0.726	0.292	0.491

Notes: Columns 1-4 represent regressions, where the left-hand-side is the average percentage squared return in the 22 days following the jump. UK data, 1930-2020H1. 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. Newey-West standard errors with 33 lags in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

Table B6: Clarity and Future Stock Volatility, UK

	(1)	(2)	(3)	(4)	(5)	(6)
Clarity	-1.421	-1.91				
	(1.56)	(1.33)				
Avg. Ease of Coding			-1.54			
			(1.38)			
Avg. Confidence				-1.7		
				(1.38)		
Share Unknown					2.01*	
					(1.14)	
Pairwise Agreement						-0.81
						(1.03)
Observations	638	638	638	638	638	638
R-squared	0.013	0.246	0.245	0.243	0.247	0.244
Return Controls	NO	YES	YES	YES	YES	YES
HAR Controls	NO	YES	YES	YES	YES	YES

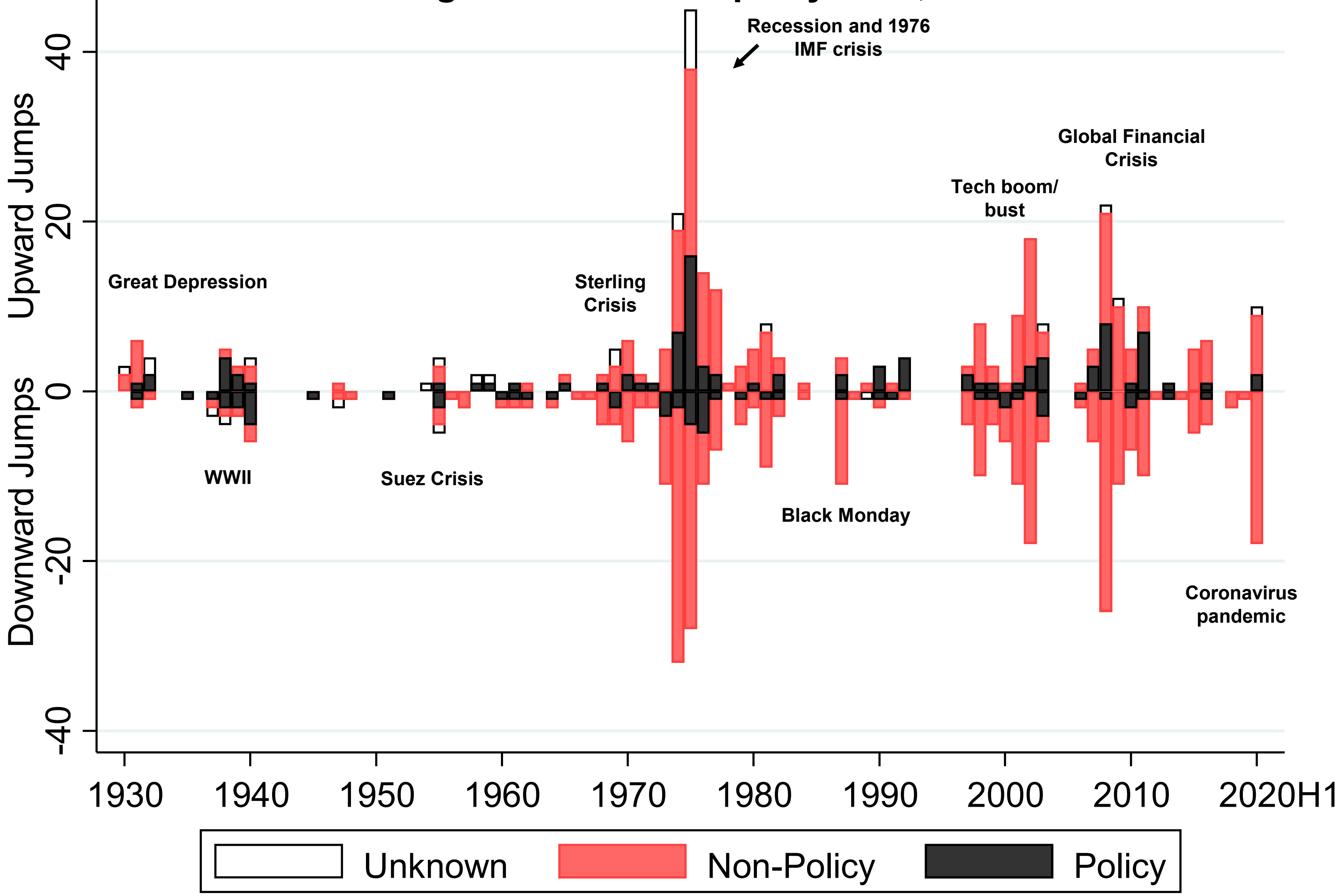
Notes: Columns 1-5 represent regressions, where the left-hand-side is the sum of squared percentage returns over the 5 days following the jump. Clarity is the first principal component of: ease of coding, confidence, share of coders who agree and share of "Unknown" codings (multiplied by negative one). It is mean zero and standard deviation one. UK data, 1930-2020H1. Robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

Table B7: Clarity and Market Concentration, UK

	Concentration x 100					
	(1)	(2)	(3)	(4)	(5)	(6)
Clarity	1.897***	1.45**				
	(0.573)	(0.627)				
Avg. Ease of Coding			0.71			
			(0.434)			
Avg. Confidence				1.67***		
				(0.559)		
Share Unknown					-0.62	
					(0.531)	
Pairwise Agreement						1
						(0.791)
Observations	166	166	166	166	166	166
R-squared	0.128	0.227	0.216	0.234	0.213	0.223
Return Controls	NO	YES	YES	YES	YES	YES
HAR Controls	NO	YES	YES	YES	YES	YES

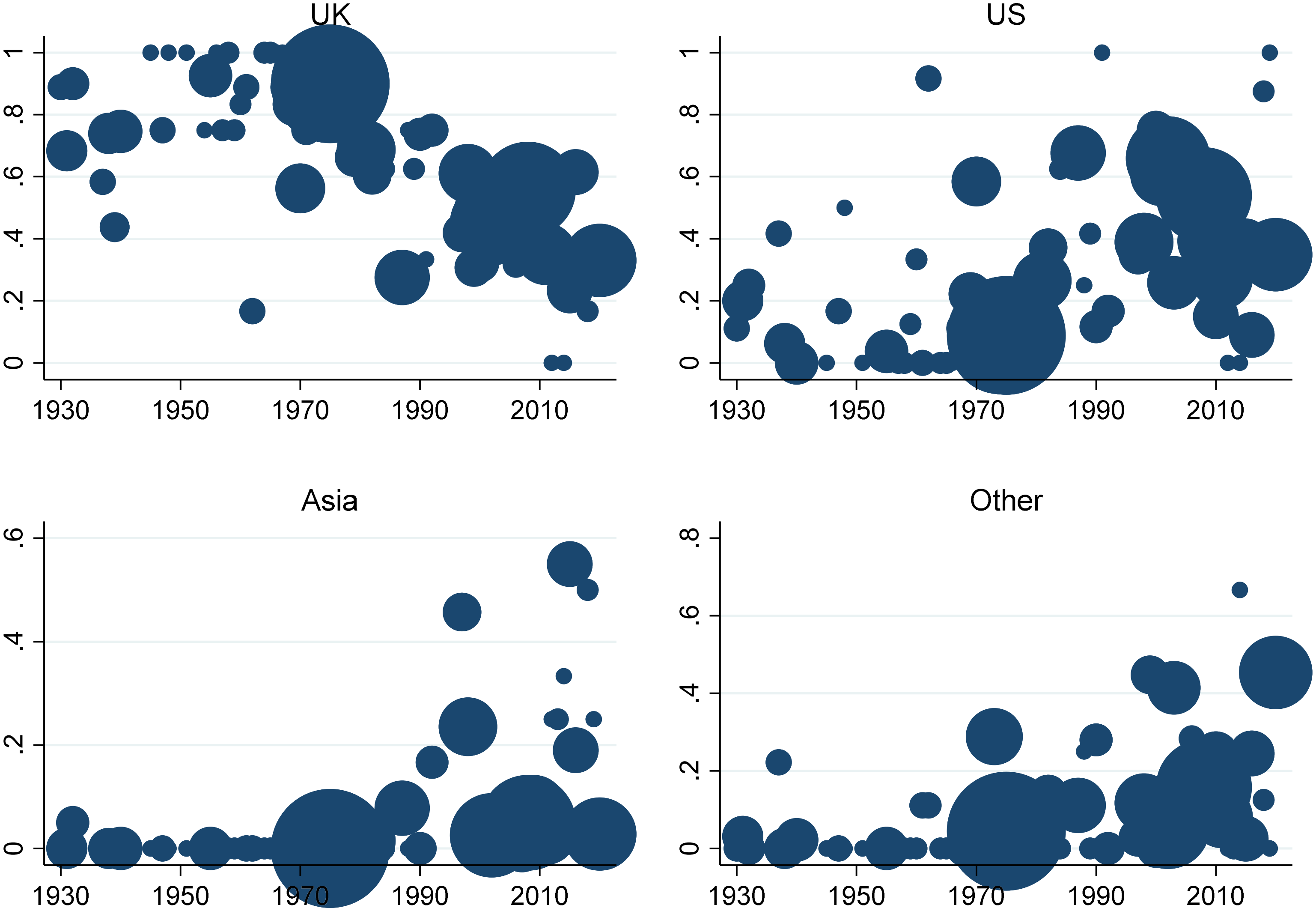
Notes: Clarity is the first principal component of: ease of coding, confidence, share of coders who agree and share of "Unknown" codings (multiplied by negative one). It is mean zero and standard deviation one. The left hand side variable is share of total distance traveled (excluding the overnight return) in the 5 minute window with the largest absolute return, multiplied by 100. Sample spans UK data for which high frequency data is available from TickWrite for the FTSE 100 Spot Market, 2006-2020H1.
*** p<0.01, ** p<0.05, * p<0.1

Figure B1: UK Jumps by Year, 1930-2020



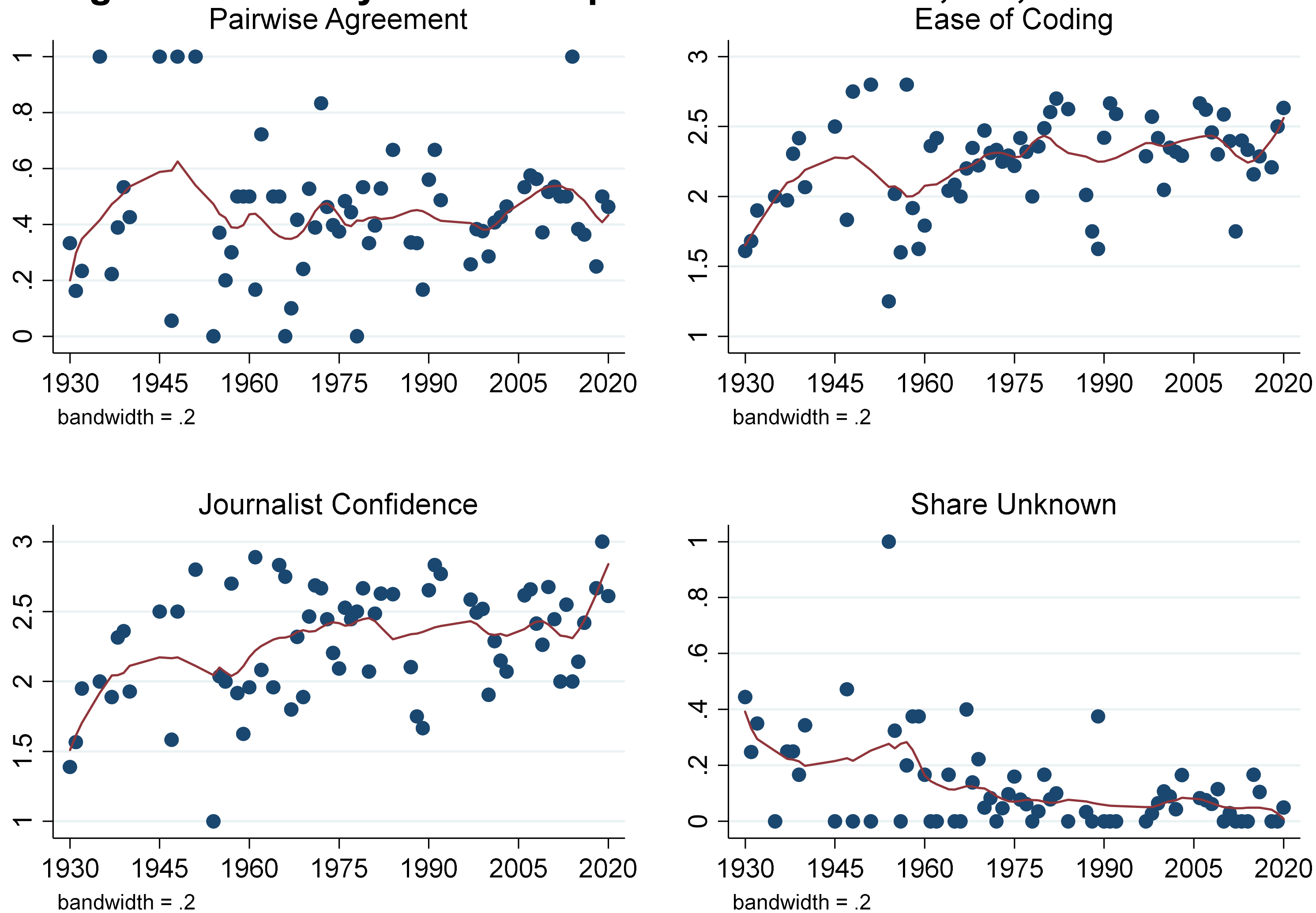
Notes: Each bar is the number of positive or negative jumps in that year. Shadings indicate the number of jumps triggered by "Policy", "Non-Policy" and "Unknown" news. Unknown includes 12 instances of "no article found". Data from 1930-2020H1.

Figure B2: Geographic Source of UK Jumps by Year, 1930-2020



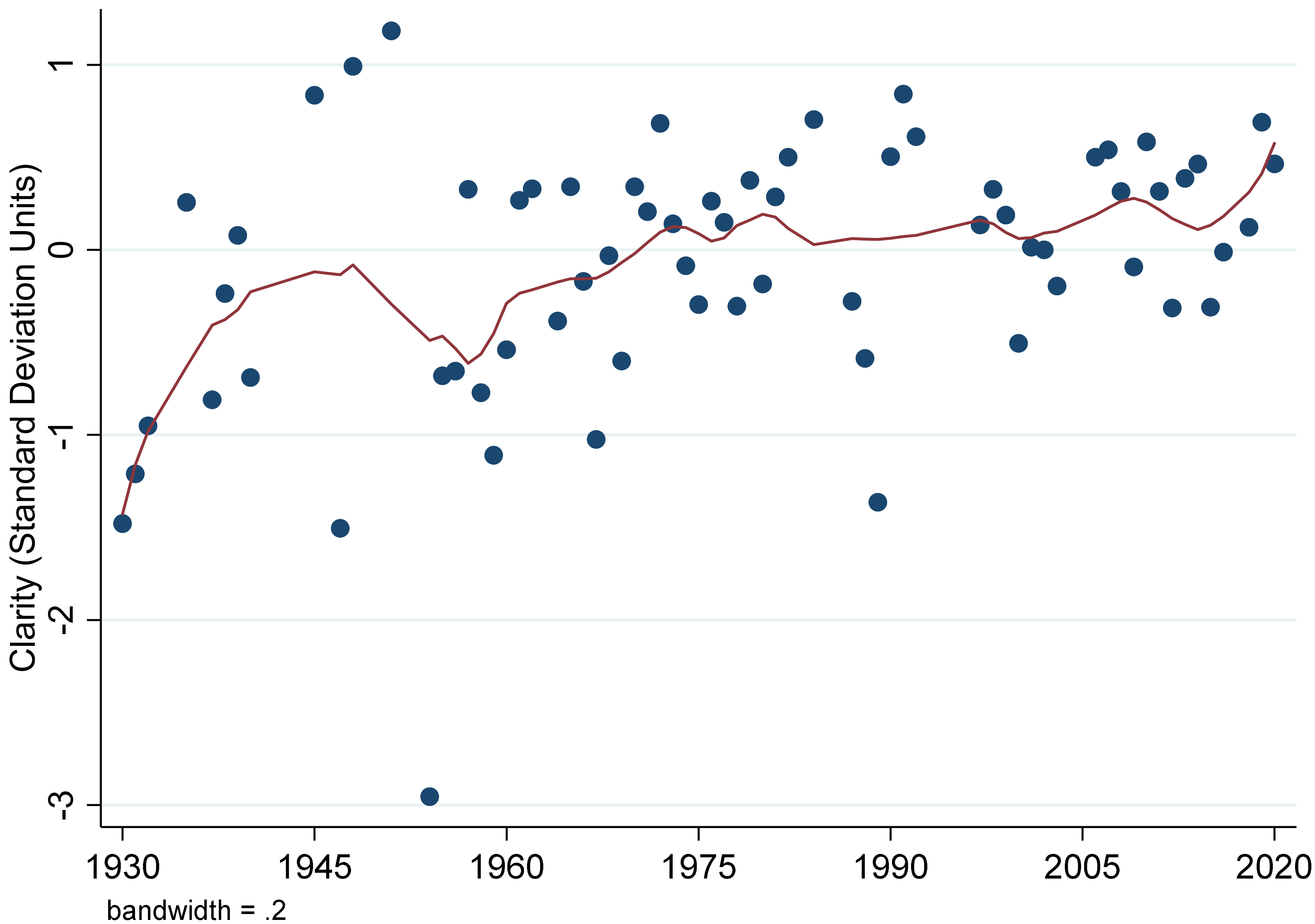
Notes: Dot shows the share of jumps in that year in the UK by their geographic origin. The size of the dots reflects the number of jumps in that year. Data from 1930 to 2020H1. Excludes unknown and no article found jumps, which have no geographic attribution.

Figure B3: Clarity Index Components Over Time, UK, 1930-2020



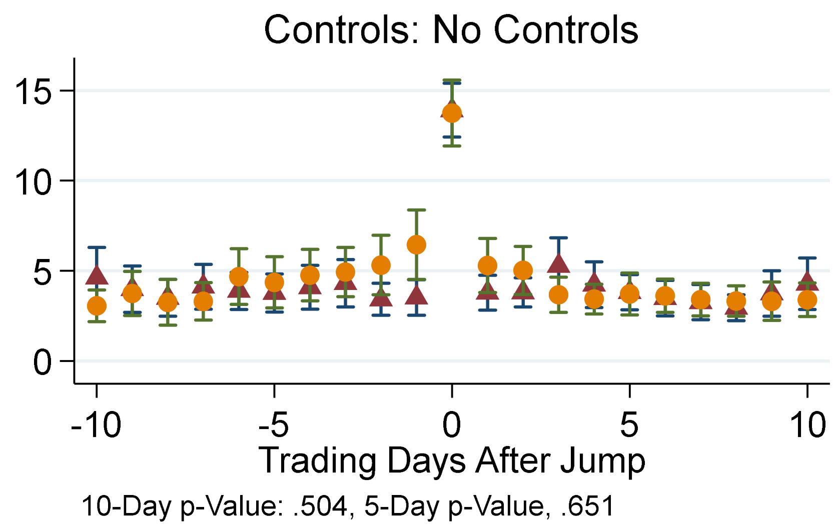
Notes: Red line represents LOWESS smoothing on components of clarity. Bandwidth is share of data used at any point in time to fit the LOWESS polynomial. Pairwise agreement is the average share of pairs of coders that agree. Ease of coding is rated on a 1-3 scale, with one being the hardest, and three being the easiest. Journalist confidence is rated on a 1-3 scale, with one being the least confident and three being the most confident. Share unknown is the percentage of coders who marked coded an article as unknown on a given day. UK data, 1930-2020H1.

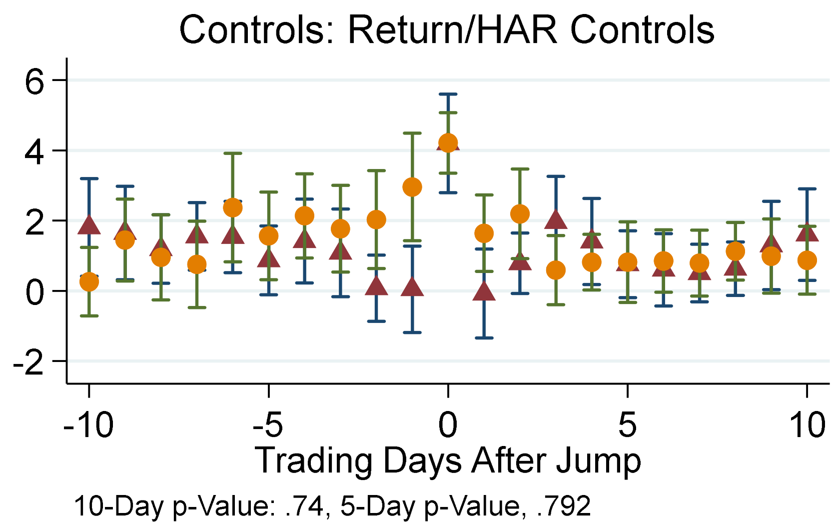
Figure B4: Clarity Index Over Time, UK, 1930-2020

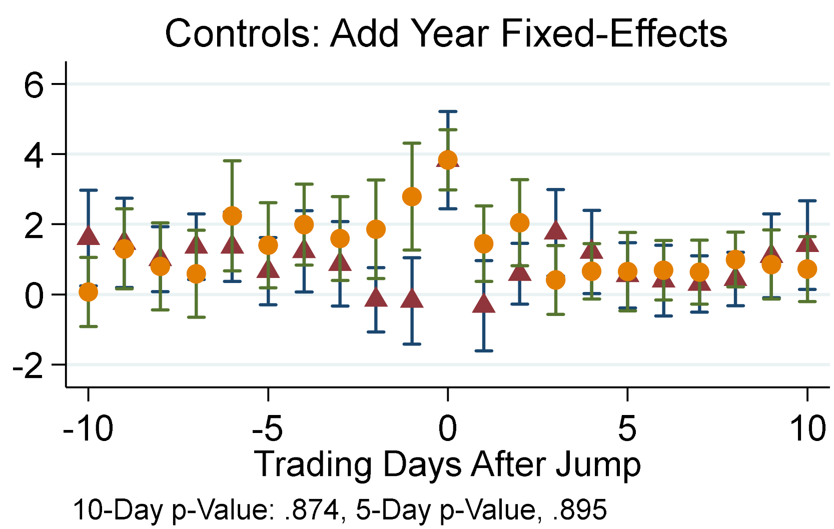


Notes: Red line represents LOWESS smoothing on clarity index. Clarity is the first principal component of: ease of coding, confidence, share of coders who agree and share of “Unknown” codings. It is mean zero and standard deviation one. Bandwidth is share of data used at any point in time to fit the LOWESS polynomial. Pairwise agreement is the average share of pairs of coders that agree. Ease of coding is rated on a 1-3 scale, with one being the hardest, and three being the easiest. Journalist confidence is rated on a 1-3 scale, with one being the least confident and three being the most confident. Share unknown is the percentage of coders who marked coded an article as unknown on a given day. UK data, 1930-2020H1.

Figure B5: Clarity and Squared Returns 10 days around jumps, UK, 1930-2020

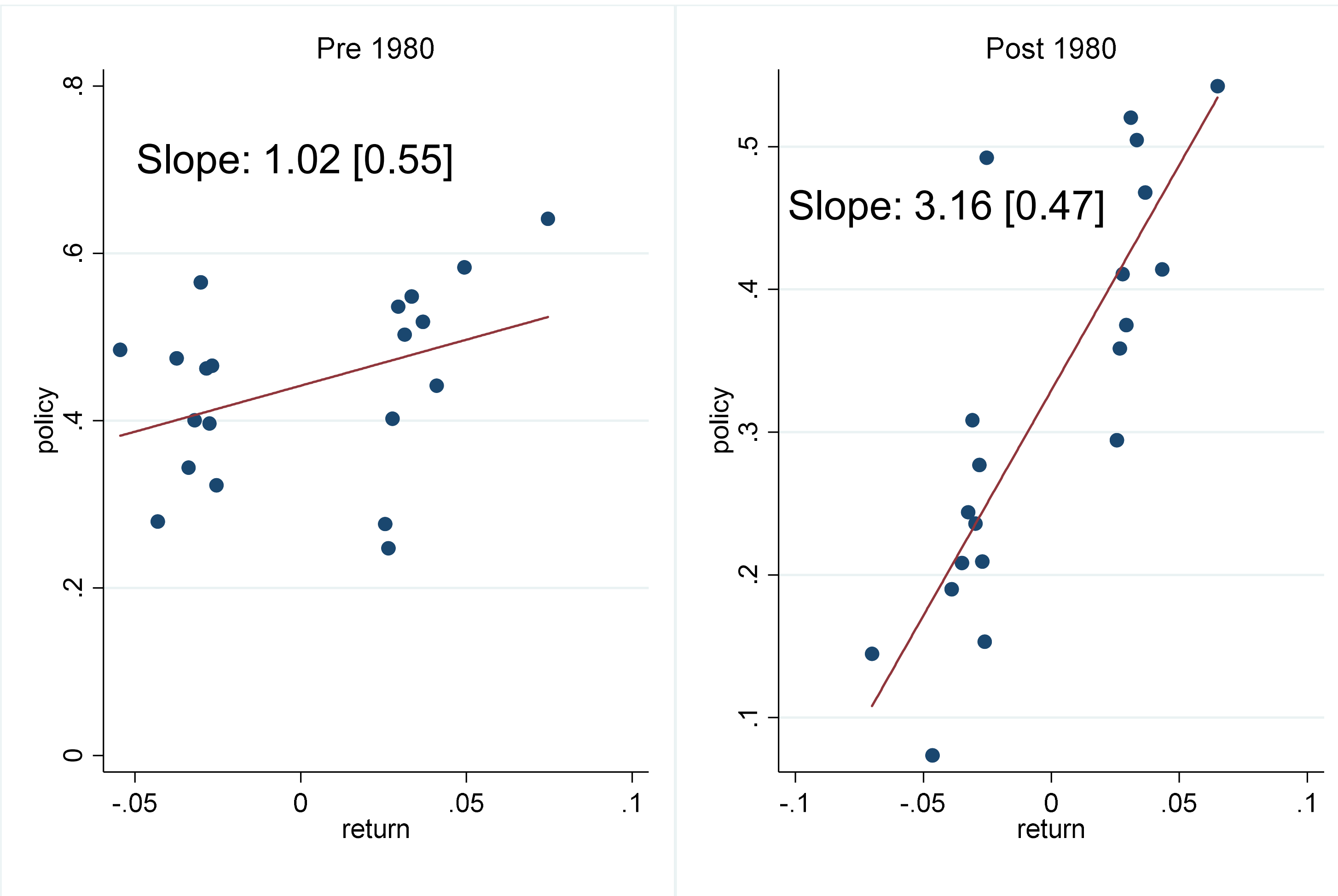






Notes: High (low) clarity is defined as clarity above (below) the sample median between 1930 and 2020. Bars around the point estimates represent 95% confidence intervals, computed with Newey-West standard errors and 10 lags. UK data, 1930-2020H1. p-Value is for joint difference between coefficients for high and low clarity jumps. Units are daily squared percentage returns.

Figure B6: Policy-Share by Jump Size and Period, UK



Realized Shares -- Diff. in slopes: 2.24, t-Stat: 9.80.
Fixing Shares -- Diff. in slopes: 1.78, t-Stat: 1.58.

Notes: Plot is a binscatter (n=20) of our policy score on stock returns. For each sub-period, we run a regression of policy on returns, and report the t-Statistic on the return variable. UK data 1930-2020H1.
We also regress (for only jump days):

$$policy_t = a + b\,return_t + c\,1_{post80} + d\,return_t \times 1_{post80} + e_t$$

And report the coefficient on the interaction term *d*, and its t-statistic at the bottom of the figure. Of the increase in slope of 2.24 we find that 20.5% is due to a shift to categories with a more positive policy mix and the remaining 79.5% is due to policy jumps becoming more positive within each category.