American Economic Review 2023, 113(6): 1395–1423
https://doi.org/10.1257/aer.113.6.1395

Presidential Address: Does Monetary Policy Matter? The Narrative Approach after 35 Years[†]

By Christina D. Romer and David H. Romer*

The narrative approach to macroeconomic identification uses qualitative sources, such as newspapers or government records, to provide information that can help establish causal relationships. This paper discusses the requirements for rigorous narrative analysis using fresh research on the impact of monetary policy as the focal application. We read the historical Minutes and Transcripts of Federal Reserve policymaking meetings to identify significant contractionary and expansionary changes in monetary policy not taken in response to current or prospective developments in real activity for the period 1946 to 2016. We find that such monetary shocks have large and significant effects on unemployment, output, and inflation in the expected directions. Analysis of available policy records suggests that a contractionary monetary shock likely occurred in 2022. Based on the empirical estimates of the effect of previous shocks, one would expect substantial negative impacts on real GDP and inflation in 2023 and 2024. (*JEL* E31, E52, E58, E65, N12)

This paper revisits one of the fundamental questions of macroeconomics: does monetary policy matter? It is a question that lies at the heart of any model of short-run macroeconomic fluctuations. If monetary policy matters, then it is vital to include a channel through which changes in aggregate demand have real effects. It is also a question that is once again at the forefront of policy discussions. The Federal Reserve is in the middle of an aggressive tightening, in hopes of lowering inflation. There appear to be widely disparate views about the likely effects of its actions on real activity, and about how quickly inflation might fall.

"Does monetary policy matter?" is also a question that is inherently hard to answer. Like so many empirical questions in economics, omitted variable bias is a central issue. Both monetary policy actions and real economic activity are likely to be influenced by other variables. Anything that affects output—expectations, fiscal policy, financial stress—is also likely to drive decisions by the Federal Reserve.

* Christina Romer: University of California, Berkeley (email: cromer@berkeley.edu); David Romer: University of California, Berkeley (email: dromer@berkeley.edu). This paper was the basis for Christina Romer's American Economic Association presidential address, presented at the ASSA Meeting in New Orleans on January 7, 2023. We are grateful to David Card, Wendy Edelberg, Emi Nakamura, Eric Swanson, Justin Wolfers, and seminar participants at the University of Cambridge for helpful comments and suggestions.

[†] Go to https://doi.org/10.1257/aer.113.6.1395 to visit the article page for additional materials and author disclosure statements.

As Kareken and Solow (1963) pointed out long ago, in the extreme case where monetary policymakers use policy to successfully counteract other forces affecting output, one would find monetary policy variables moving all around and output not changing. A simple regression of output on an indicator of monetary policy would naïvely and incorrectly conclude that monetary policy didn't matter. That is, if countercyclical monetary policy actions are common, the estimated impact of monetary policy will be biased toward zero.

In 1989, we wrote a paper on the effects of monetary policy using a technique we termed the "narrative approach" (Romer and Romer 1989). This is an empirical technique where one gathers systematic evidence from contemporaneous qualitative sources (such as newspapers, government reports, and policy meeting transcripts), and incorporates it into statistical analysis. In the case of our 1989 study, the idea was to deal with omitted variable bias by using the plentiful narrative record on the motivation for monetary policy actions. We used the records of the Federal Reserve to identify a subset of monetary actions that were not motivated by other factors affecting output. We argued that the behavior of output following these shocks would provide relatively unbiased estimates of the impact of monetary policy.

The subtitle of that paper, "A New Test in the Spirit of Friedman and Schwartz," captured our intellectual debt to Milton Friedman and Anna Jacobson Schwartz. Friedman and Schwartz pioneered the narrative approach in their seminal work, *A Monetary History of the United States* (1963). We like to think that we have built on their insights and added some modern rigor to the approach.

Over our careers, we have used the narrative approach in a number of studies. For example, we used it to separate tax changes into those taken for countercyclical reasons or because spending was increasing, and those taken for more exogenous reasons (Romer and Romer 2010). We found that the estimated response of output to tax changes was faster and larger when we limited the analysis to exogenous changes.

Others have also used the narrative approach to achieve macroeconomic identification. James Hamilton used it to argue that many large oil price increases resulted from factors outside the US macroeconomy (Hamilton 1985). This suggested that the correlation he found between oil price shocks and recessions reflected a largely causal relationship (Hamilton 1983). Valerie Ramey used narrative sources to identify when news of changes in defense spending became known (Ramey 2011). This information allowed her to better estimate the size and timing of the effects of government spending. Andrew Jalil used the narrative approach to identify prewar financial crises in the United States, and to isolate a subset that were due to factors uncorrelated with macroeconomic conditions—such as managerial malfeasance or other idiosyncratic shocks (Jalil 2015). Incorporating this additional information allowed him to identify the true causal impact of financial crises.

This address is an opportunity to reflect on the narrative approach, and to revisit our first paper using it. The heart of the address involves redoing our 1989 study. We feel we have learned a great deal about the pitfalls of narrative research and how to do it better over the past three-plus decades. In Section I, we discuss those lessons, and apply them to the identification of monetary policy shocks in the United States since 1946.

Having redone the narrative work, we then incorporate the new evidence into a statistical framework. In Section II, we discuss both our methodology and our

findings about the impact of monetary policy on real economic indicators, such as the unemployment rate and real GDP. We find that a contractionary monetary policy shock increases the unemployment rate by 1.6 percentage points and reduces real GDP by 4.4 percent, relative to the no-shock baseline. The effects of monetary policy are highly statistically significant. We find that our improved series on monetary policy shocks results in slightly smaller estimates of the real effects than we found in our 1989 study, but greater statistical precision. We also examine the impact of monetary shocks on inflation. We find that contractionary monetary policy reduces inflation by about 1.5 percentage points, but the effects develop slowly and are less precisely estimated than those for real activity.

Though our main focus is on the narrative approach and estimates of the impact of monetary policy, in Section III we consider the implications of our findings for current monetary policy. What does our new and improved narrative work and empirical estimation tell us about the likely outcome of the Federal Reserve's recent actions to contain inflation?

Finally, in Section IV, we consider the future of the narrative approach. Is it still needed, or has it been superseded by other approaches? And, if it is still needed, does it need to be done by human researchers?

I. Rigorous Narrative Analysis

The key step in narrative research is to analyze some narrative source to gather systematic evidence that can be used in an empirical framework along with conventional data. Many years ago, someone referred to what we do as the "literary" approach. It was clear they didn't mean it as a compliment. The tone was that narrative work is inherently less scientific and rigorous than conventional data analysis. We disagree strongly. We believe deeply that narrative evidence can be both rigorous and reproducible.

A. *Features of Good Narrative Analysis*

In this section, we describe what we have learned are some of the requirements for rigorous narrative analysis. These requirements are summarized in Table 1. We illustrate the features with the new narrative research we have been doing on the effects of monetary policy—with an eye toward showing how our work has changed in light of what we have learned over the past 35 years.

A Reliable Source.—A fundamental determinant of the quality of any narrative evidence is the reliability of the source. The source should be contemporaneous or real time. Otherwise, subsequent economic outcomes could affect how the source describes or interprets actions. The narrative source should exist for a long period of time and be reasonably consistent over time. Otherwise, one is not getting the same information over the full sample. The source needs to be accurate. One is looking for a narrative source that accurately reports what was said at a meeting, or the observations of knowledgeable, accurate observers.

For our current study, we are using the historical *Minutes* and *Transcripts* of the meetings of the Federal Open Market Committee (FOMC) of the Federal Reserve

Table 1—Requirements for Rigorous Narrative Analysis

1. A reliable narrative source
- Real time
- Consistent over time
- Detailed and accurate

2. A clear idea of what one is looking for in the source
- Specify criteria in detail

3. Approach the source dispassionately and consistently
- Resist the temptation to see what you want to see in the source
- Compare classifications with another reader
- Read from beginning to end
- Don't use previous knowledge to focus on certain periods

4. Document the narrative evidence carefully
- Force yourself to explain your reasoning
- Make it easy for others to check your work

System.[1] These are either very detailed summaries of the discussion with extensive paraphrases, or verbatim transcripts. They are contemporaneous with the monetary policy decisions. There are roughly 50 to 100 pages of material per meeting—so, with eight to twelve meetings per year (or in some periods even more), we are talking about a lot of information (and reading!). We start in 1946 because we are interested in the effects of monetary policy in the post-World War II period.

One important virtue of this source is that, for much of our sample, it was confidential. The detailed records of FOMC meetings were not intended or expected to be released to the public. As a result, there was little reason for policymakers not to speak honestly and unguardedly. Starting in 1993, the *Transcripts* were made public, but with a five-year lag. The lag likely helped to keep the discussions frank.[2] It also means that our analysis has to end in 2016, which is the latest year for which the *Transcripts* have been released.

The source we are using is different from that in our original study. Back in 1989, we couldn't face the thousands of pages of historical *Minutes* and *Transcripts*. So we instead mainly used the much briefer "Record of Policy Actions of the Federal Open Market Committee." These "Records" are very similar to what are now called the "Minutes of the Federal Open Market Committee." They are quite brief—on the order of just a few pages per meeting early in the postwar period to about ten pages per meeting in the 1980s. Because these "Records" are much less detailed, they inevitably provide less information on the motivation for policy. Moreover, because the "Record of Policy Actions" was always made public with only a short lag, we suspect it may be less forthright. For these reasons, our current narrative source has substantial benefits over the earlier one.

[1] The narrative sources are described in more detail in online Appendix A. They are all available on the Board of Governors website: https://www.federalreserve.gov/monetarypolicy/fomc_historical_year.htm.

[2] Meade and Stasavage (2008) and Hansen, McMahon, and Prat (2018) find some changes in the meetings after participants learned that transcripts would eventually be released, but no clear changes in the frankness of the discussion.

A Clear Sense of What One Is Looking for in the Narrative Source.—Just as important as the narrative source is how one approaches it. The key lesson is that you need to have a clear idea of what you are looking for in the documents. This is where the creativity lies. What can we learn from the narrative source that would aid in establishing causation? This is what made Friedman and Schwartz's *Monetary History* so important. Their genius was in knowing what to look for in the diaries of monetary policymakers and other early Federal Reserve records.

What are we looking for in the narrative record for this revisiting of the effects of monetary policy? At a very broad level, we are looking for times when monetary policymakers changed money growth and interest rates for reasons unrelated to current or prospective real economic activity. These are policy "shocks" in the sense that monetary policy is not being driven by output or other factors affecting output. Thus, the behavior of output and other indicators of real activity following such episodes should give relatively unbiased estimates of the causal impact of monetary policy.[3]

The problem with this broad framing is that it is hard to specify concrete criteria that can be used to determine when such episodes occurred. So instead, we look for something more specific in the narrative record. As in our original paper, we look for times when monetary policymakers felt the economy was roughly at potential (or normal) output, but decided that the prevailing rate of inflation was too high. Policymakers then chose to cut money growth and raise interest rates, realizing that there would be (or at least could be) substantial negative consequences for aggregate output and unemployment. These criteria are designed to pick out times when policymakers essentially changed their tastes about the acceptable level of inflation. They weren't just responding to anticipated movements in the real economy and inflation.

The requirements that policymakers were taking actions and were willing to accept real economic consequences are designed to ensure that policymakers were not just engaged in ritual hand-wringing. In terms of the simple aggregate supply/aggregate demand framework, we are looking for times when the Federal Reserve deliberately shifted the aggregate demand curve back. If monetary policy has real effects, output should fall after such actions. That is, these are contractionary monetary policy shocks.

Focusing on a very specific type of policy movement unrelated to current or prospective economic developments comes with both a cost and a benefit. The cost is that it means we will not be using some episodes that are legitimate for estimating the causal output effects of monetary policy. This would likely decrease the precision of our estimates, but should not introduce bias. The benefit is that makes our criteria and identification procedures much more concrete, and so less prone to errors.

[3] Importantly, we use the term "shock" to mean movements in monetary policy unrelated to current or prospective real economic activity. They needn't be unanticipated when they occur. That is, if the Federal Reserve telegraphs a change in policy, it can still be used to estimate the effects of policy as long as it is not driven by current or prospective real economic activity. Similarly, the change in policy needn't be a departure from the Federal Reserve's usual behavior. If monetary policymakers respond in their usual way to something unrelated to current or prospective real output, that still constitutes a shock by our definition.

An important extension that we do in the new work is to broaden the criteria to include expansionary monetary policy shocks. In particular, we now also look for times when policymakers believed that they were at a stable level of economic activity, but took actions to lower the unemployment rate—and were willing to accept adverse consequences for inflation. That is, we look for times when policymakers were deliberately shifting the aggregate demand curve out because of a change in their view of the acceptable or desirable level of unemployment. If monetary policy has real effects, output should rise following such actions.

Approach the Narrative Source Dispassionately and Consistently.—To do the narrative analysis rigorously, one needs to approach the qualitative source dispassionately and consistently. Researchers need to resist the temptation to see what they want to see in the narrative source. We find that having multiple readers for the same document provides a valuable counterbalance to this temptation. Likewise, one has to read the entire source from beginning to end—not look harder for information in some periods than others.

Document the Narrative Evidence Carefully.—We have also found that writing detailed descriptions of our analysis of the narrative evidence is a crucial part of the work. The documentation process clarifies our own thinking and sometimes makes us change our minds about the classification of an episode. And, it makes it easier for subsequent researchers to check our work. One change that we have noticed in our research is that these narrative appendices have been getting longer and longer.

B. *Examples of the Narrative Analysis of Monetary Shocks*

To provide a feel for how we apply our criteria, here we describe some of the evidence for two episodes we identify as monetary shocks. The much more detailed description of the narrative evidence for these two shocks, as well as the others we identify, is contained in online Appendix A.

Contractionary Shock in December 1988.—December 1988 is an example of a contractionary monetary policy shock. It is the latest contractionary shock in our sample (which, as noted above, ends in 2016).

For much of 1987 and 1988, policymakers were worried that inflation would pick up if they didn't tighten. In response, they made small moves toward restraint. These initial moves don't count as a contractionary shock by our criteria. Policymakers weren't trying to shift the aggregate demand curve back from a stable level; rather they were just trying to hold it steady against other forces tending to shift it out.

In May 1988, policymakers started to say that the current level of inflation was unacceptable, not just that they wanted to prevent inflation from rising further. For example, one said: "In terms of our own inflation rate … we have been stalled at a rate that I think is too high for most of us" (W. Lee Hoskins, *Transcript*, May 17, 1988, p. 5).[4] Another said: "whatever is likely to happen on the wage and price side,

[4] All citations to the *Transcripts* are to US Board of Governors of the Federal Reserve System (1976–2016).

it doesn't seem to me that there's going to be any deceleration next year unless we act. I think it is time for some further action" (Gary Stern, pp. 4–5). However, a number of other members were less clear that they wanted to literally reduce inflation and were hesitant to incur the possible output consequences. A concern about rising inflation, rather than its level, continued through the summer and fall of 1988. For example, at the August 1988 meeting, one member said, "the bottom line is that we are in a territory of accelerating inflation and we have to resist that growth" (Edward Boehne, *Transcript*, August 16, 1988, p. 17).

A desire to reduce inflation and a willingness to accept output consequences became much more widespread at the December 1988 meeting. For example, one member said: "I think the job before us is to contain the inflation and to slow this economy down" (Forrestal, *Transcript*, December 13–14, 1988, p. 56). Another said that, while the fragility of the financial system was a real concern, "we can't design monetary policy to avoid any difficulties in various sectors. …[W]e've got to focus on inflation and if something goes wrong then you can address those problems" (H. Robert Heller, p. 51). The staff predicated their forecast on a goal of reducing inflation. They warned that "if it is the aim of the Committee not merely to hold the line on inflation but, rather, to restore a downward trend by 1990, then it may be necessary to run the risk of some financial stress and economic weakness" (Michael Prell, *Transcript*, Presentation Materials, December 13–14, 1988, p. 1). FOMC members nevertheless agreed to embark on a significant tightening.

This episode counts as a contractionary monetary policy shock because, at a stable level of growth and unemployment, policymakers decided that the current level of inflation was unacceptable and took actions to reduce it. And they clearly understood and accepted that there could be substantial adverse consequences for output and unemployment.

Expansionary Shock in January 1972.—January 1972 is an example of an expansionary monetary policy shock. Indeed, it is the only expansionary shock we find.

Following the recovery from the mild recession of 1969–1970, there was a sense among FOMC members and the staff that unemployment had stabilized at an elevated level. For example, in November 1971, the Federal Reserve staff projected that unemployment would fall only slightly (from 6 percent to 5.3 percent) in 1972 (Joseph Zeisel, *Memorandum of Discussion*, November 16, 1971, p. 40).[5]

By December 1971, there was a growing consensus that the prevailing level of unemployment was unacceptably high, and that the Federal Reserve needed to take expansionary actions. For example, one member "believed the appropriate posture for the System at this point was one of doing what it could with the policy instruments at its disposal to foster and encourage economic expansion" (J. Dewey Daane, *Memorandum of Discussion*, December 14, 1971, p. 60).

[5] All citations to the *Memoranda of Discussion* are to US Board of Governors of the Federal Reserve System (1967–1976). In 1967, following the passage of the Freedom of Information Act, the Federal Reserve separated out a small amount of the material in the historical *Minutes* and called them the "Minutes of Actions." The remaining material, which appears to contain essentially all of what was previously in the historical *Minutes*, was put in a new document called the *Memorandum of Discussion*. Like the historical *Minutes*, the *Memoranda of Discussion* are written in the third person and contain extensive paraphrases of the discussion.

In a move very similar to one used by Federal Reserve Chairman Paul Volcker in 1979, Chairman Arthur Burns called a special meeting of the FOMC on January 11, 1972 "because he had become seriously concerned about the present stance of monetary policy" (*Memorandum of Discussion*, January 11, 1972, p. 4). He said that "unless the aggregates now began to grow at adequate rates, he would … feel that there might be some validity in a charge that the System was not supporting the policies of the Administration and Congress" (p. 62). In another parallel to Volcker's actions in 1979, Burns also put on the agenda a proposal to adopt reserve targeting. And, he was very clear that "there could be a further reduction in interest rates, possibly of significant dimensions" (pp. 63–64).

A number of FOMC members feared that the switch to rapid reserve growth and lower interest rates would generate inflation. For example, one said that "the Committee should consider whether stimulating the economy to greater heights in the short run would not involve a cost in the form of a resurgence of inflationary pressures later on" (Philip Coldwell, *Memorandum of Discussion*, January 11, 1972, p. 71). Another said that "it should be recognized that the battle against inflation was not yet over, and that unduly aggressive policy actions would involve the risk of rekindling inflationary expectations" (James Robertson, p. 90). Nevertheless, despite three dissenting votes, a majority of the committee agreed to radically more expansionary policy.

This episode meets our criteria for an expansionary monetary policy shock. From a stable level, the FOMC decided to take action to lower unemployment because it felt the current level was unacceptable. Policymakers were deliberately shifting out the aggregate demand curve, despite widespread concern that it would be inflationary.

C. *Results of the Narrative Analysis*

The outcome of all this reading and analysis is the identification of ten dates of monetary policy shocks over the period 1946–2016. Nine are contractionary shocks and one is expansionary. These dates are shown in Table 2.

Interestingly, although we expanded our sample by almost 30 years, we found no monetary shocks between 1988 and 2016. At some level, this is not terribly surprising. Inflation, until very recently, had been low and steady for almost 30 years. And, American monetary policy has been widely viewed as deft and carefully calibrated. It was to be expected that there might be nothing to call a clear-cut anti-inflationary monetary policy shock since the late 1980s. Obviously, however, that may have changed in the past year—when inflation once again rose substantially and monetary policymakers have taken strong actions to reduce it. We discuss the recent experience in Section III.

Part of the motivation for revisiting this topic and narrative work was to see how well we had done originally. This is a test both of our younger selves and our more limited narrative source. For the most part, we were pretty happy with our original findings. In addition to the new dates of monetary shocks, Table 2 also shows the dates we set originally.[6] We agreed that all our original dates were indeed contractionary

[6] The December 1988 shock included as an "original" shock was not in Romer and Romer (1989). We identified it in Romer and Romer (1994).

TABLE 2—MONETARY POLICY SHOCKS, 1946–2016

New dates			Original dates		
October	1947	(−)	October	1947	(−)
August	1955	(−)	September	1955	(−)
September	1958	(−)			
December	1968	(−)	December	1968	(−)
January	1972	(+)			
April	1974	(−)	April	1974	(−)
August	1978	(−)	August	1978	(−)
October	1979	(−)	October	1979	(−)
May	1981	(−)			
December	1988	(−)	December	1988	(−)

Notes: Contractionary shocks are denoted (−) and expansionary shocks are denoted (+). In setting our original dates, we did not have a classification for expansionary shocks.

monetary shocks. Looking at more detailed records caused us to change the timing of the shock by a month in one case, but no major changes.

We did, however, find two contractionary shocks that we had missed in our original study. One is in September 1958, when monetary policymakers switched to contractionary policy very quickly after the 1957 recession ended because they were concerned about the current level of inflation. It is quite clear in the detailed historical *Minutes*, but even knowing that, we cannot see it in the brief, public summaries we used previously.

The other new contractionary shock is in May 1981, and reflects a more nuanced reading of the Volcker era. There is no doubt that there was a contractionary monetary policy shock in October 1979, when the Federal Reserve embarked on what has come to be called the Volcker disinflation. But following the severe downturn in the second quarter of 1980, we detect in the *Transcripts* a definite change in focus. Despite continuing high inflation, there was a sustained period where there was little interest in aggressive inflation reduction and where policy was dramatically looser. But then in late 1980 and early 1981, there was a gradual shift back to widespread agreement among FOMC members that the current level of inflation was unacceptable, and they were willing to risk a recession to deal with it. The funds rate was allowed to rise dramatically again. Because this change came after the FOMC had moved away from inflation reduction, we now think it should be classified as a separate shock.

Finally, as already described, broadening our criteria to allow for expansionary monetary policy shocks revealed one such episode.

D. *Comparing Interpretations of Various Episodes*

Because of the importance of relying on contemporaneous sources in narrative work, we did not use retrospective analyses in our identification of monetary shocks. But of course, numerous other researchers have also examined various episodes of postwar monetary policy. Though these other scholars were often looking for different things in the narrative evidence, we find that they confirm some important conclusions of our analysis.

For example, we are far from the first to find that the move to anti-inflationary monetary policy under Paul Volcker was more complicated than just a one-time switch in October 1979. Goodfriend and King (2005) conclude, "The Volcker Fed's initial inflation-fighting effort was abandoned in mid-1980" (p. 986), and, "The true onset of the Volcker disinflation dates to November 1980 or early 1981" (p.1000).[7] In discussing policy from roughly March through July 1980, Meltzer (2009, p. 1056) says, "the Federal Reserve responded to high interest rates and rising unemployment by giving up its anti-inflation policy." He is not precise about when he sees a shift back to "anti-inflation policy" (p. 1085), but he puts the most emphasis on Spring 1981: "policy tightened sharply in the spring of 1981. By increasing interest rates in the middle of a recession, the FOMC increased its credibility. … Never before had the public seen an increase in interest rates with the unemployment rate at 7.5 percent" (pp. 1081–82).

As another example, Friedman and Schwartz's interpretation of the Federal Reserve's tightening in the late 1950s is very similar to ours. They refer to a "sharp reversal in monetary policy … beginning in 1959" (while we date the shift in September 1958), and they see a "renewed emphasis by the Federal Reserve System upon the dangers of inflation rather than of contraction" (Friedman and Schwartz 1963, p. 617). And, very much in line with what we describe in online Appendix A, they say, "retrospective examination of its earlier policy persuaded the Reserve System that it had erred during the 1954–57 expansion by continuing to 'ease' for too long …. It was determined not to repeat the error" (pp. 617–18).

At the same time, two episodes where we identify shocks have received relatively little attention from others. The first is 1947. Both Friedman and Schwartz (1963, pp. 577–80), who did not have access to the detailed *Minutes* for this period, and Hetzel (2008, pp. 37–40) discuss monetary policy in 1947 only briefly, and in the context of larger discussions of monetary policy in the postwar period before the 1951 Federal Reserve–Treasury Accord. Meltzer (2003) devotes more attention to this episode—including describing the October 1947 FOMC meeting as a "turning point" (p. 652)—but does not characterize it. The other is 1978. Hetzel (2008) barely mentions the policy tightening in this period. Meltzer (2009, pp. 925–39) describes the tightening, but does not characterize it as an important anti-inflationary policy shift. Nelson (2005) is something of an exception, seeing the Federal Reserve as shifting to tighter policy in an effort to reduce inflation, but largely motivated by a belief that the exchange rate appreciation caused by higher interest rates would act as a favorable "cost-push" shock.

We view the fact that we identify monetary shocks in some episodes that have received little attention from others as a desirable consequence of our efforts to implement the narrative approach in ways that minimize bias. We find the Federal Reserve's anti-inflationary motivation and actions around August 1978 to be almost as strong as in our other episodes. We suspect that the combination of inflation's failure to fall, the fact that a recession did not materialize in 1979 (despite being

[7] Although we date the shock as occurring in May 1981, we see the beginning of a renewed commitment to bringing inflation down in December 1980. The earliest evidence Goodfriend and King provide of a shift in the Federal Reserve's intentions comes from the transcript of the February 1981 FOMC meeting. Thus we are in broad agreement about the timing of the change.

widely expected at the time), and the more dramatic events of October 1979 have drawn attention away from the 1978 episode. But if the goal is to estimate the effects of monetary policy, these are not grounds for excluding it. And, we suspect the 1947 episode may be largely overlooked because it predates the Accord. But as we describe in online Appendix A, the Federal Reserve had a strong desire to bring inflation down, took actions it viewed as important, and had tools it thought were sufficient for the task (including the ability to move short-term interest rates). We view our identification of this shock as reinforcing the case for reading the entire narrative record uniformly.

II. Incorporating Narrative Evidence into a Statistical Framework

Having done the narrative work, the crucial next step is to incorporate the narrative evidence into a statistical framework. As mentioned earlier, Friedman and Schwartz were the pioneers of rigorous narrative analysis. Their reading of Federal Reserve records and the diaries of monetary policymakers is second to none. But a weakness of their work was a reliance on somewhat informal empiricism. Having identified some crucial episodes where monetary policy was contractionary for reasons unrelated to the state of the economy, they didn't do much beyond observing that output and prices fell afterward. Narrative evidence is much more valuable if it is analyzed using appropriate statistical techniques.

A. *Methodology*

For our revisiting of the impact of monetary policy, we follow a very straightforward approach. We use the dates of monetary policy shocks to create a dummy variable. We set it equal to 1 in the months (or quarters) of a contractionary shock, -1 in the month (or quarter) of an expansionary shock, and zero otherwise. We then regress the outcome variable of interest on the dummy variable.

More specifically, we use a Jordà local projection approach (Jordà 2005). The Jordà approach involves running a series of regressions of some outcome variable at various horizons (h) after time t on the independent variable of interest at t and control variables. In our case, we estimate the following sequence of equations for different values of h:

$$Y_{t+h} = \alpha^h + \beta^h S_t + \sum_{k=1}^{K} \varphi_k^h S_{t-k} + \sum_{k=1}^{K} \theta_k^h Y_{t-k} + e_t^h, \tag{1}$$

where Y is the outcome variable of interest and S is the dummy variable for the dates of monetary shocks. The sequence of the estimated βs for the various horizons is an estimate of the response of the outcome variable to a realization of 1 for our dummy variable. Given our sign convention, it thus traces out the impulse response function to a *contractionary* monetary policy shock.

We include as control variables lags of both the outcome variable and our shock dummy. K is the number of lags included. For monthly outcome variables we include twelve lags of both series; for quarterly outcome variables, we include four lags of each. The inclusion of lags of the outcome variable captures the usual dynamics of the series. That way the shock dummy picks up the effects of a monetary shock

relative to the normal behavior of the outcome variable. We estimate this specification for each horizon from 0 to 60 months (or 0 to 20 quarters) ahead. We use data for 1946:10 to 2016:12 (or 1946:IV to 2016:IV for quarterly series).[8]

We consider a range of outcome variables. For real economic activity, we use the monthly unemployment rate and quarterly real GDP. We also consider quarterly inflation, measured using the GDP price index, the price index for personal consumption expenditures (PCE), and the price index for PCE excluding food and energy. The sources for the data are given in the Data Appendix.

Now there are obviously other empirical procedures one could use. For example, instead of a simple dummy variable for the dates of monetary policy shocks, one could try to scale the dummy in some way—perhaps by the change in interest rates over the episode or by using the same narrative sources to classify the severity of the shock. Though we have not attempted to use the narrative sources to scale the shocks, the evidence discussed in online Appendix A does not suggest large variations in their severity.

In place of a reduced-form regression, one could use the shock dummy as an instrument in a regression of the outcome variable of interest on some quantitative measure of monetary policy. We have not followed this path because the nature of monetary policy has changed substantially over time. While the Federal Reserve has primarily used interest rate changes as the main tool of monetary policy since the mid-1980s, in the 1950s and 1960s it was much more common for it to use changes in reserve requirements and credit restrictions as well. As a result, it is hard to find a single quantitative indicator appropriate for the last 70 years.[9] But in other applications, an IV approach can be a very sensible way to incorporate narrative evidence.

B. *Results*

Unemployment Rate.—Figure 1 shows the estimated impulse response function of the monthly unemployment rate to a contractionary monetary policy shock, along with the two-standard-error bands. Because the unemployment rate is entered in the regression in levels, the estimated impulse response function shows the impact of a monetary policy shock on the unemployment rate in percentage points relative to the no-shock baseline.

In response to a contractionary monetary policy shock, the unemployment rate rises gradually—starting about 5 months after the shock. The maximum impact is a rise of 1.6 percentage points after 27 months. The peak effect is highly statistically significant, with a t-statistic of 3.5. This effect is fairly large. In the recessions in our sample, unemployment rose on average by about 2.8 percentage points. Thus, our estimates imply that a monetary shock causes a rise in unemployment more than

[8] Since we have twelve lags and the dependent variable at horizon h is Y_{t+h}, this implies that the sample period for the horizon h regression with monthly data is 1947:10 to h months before 2016:12. We use data starting in 1946:10 so that even with twelve lags we are able to include the first nonzero value of our shock variable (which is in 1947:10). We stop in 2016:12 because that is when our shock series ends. With quarterly data, the sample is 1947:IV to h quarters before 2016:IV.

[9] The fact that early monetary policy actions were a mix of interest rate changes and quantitative restrictions also means that scaling the shocks by the interest rate change would give a misleading sense of the relative severity of the shocks.

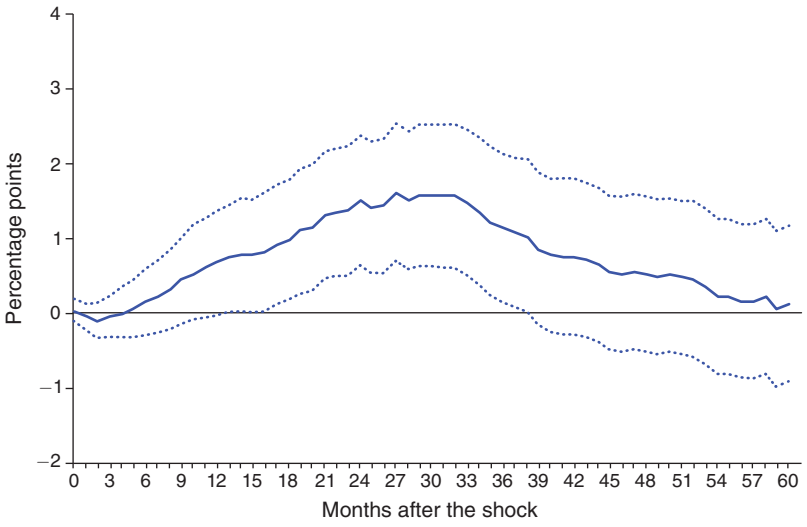

FIGURE 1. RESPONSE OF THE UNEMPLOYMENT RATE TO A MONETARY POLICY SHOCK

Notes: The figure shows the results of estimating equation (1) for horizons 0 to 60. The dependent variable is the unemployment rate. The dotted lines show the two-standard-error confidence bands. The new shock series is given in Table 2. See text for details of the estimation and the Data Appendix for the sources of the unemployment series.

half as large as a typical recession. The effect of a monetary policy shock on unemployment goes away almost entirely by five years after the shock.

One thing we can do is compare the results using our new and improved shock series with those using our original dates. This provides a read on whether the new dates matter. Figure 2 shows the two impulse response functions. The results using the new series are in blue, those using the original dates are in red. The point estimates are slightly smaller with the revised dates, but are estimated more precisely.

Real GDP.—Real GDP, which is a quarterly series, is entered in the regressions in logarithms (times 100). The sequence of coefficients on the shock dummy for various horizons shows the impact of a contractionary monetary policy shock, measured as the percentage deviation from the no-shock baseline.

Figure 3 shows the estimated impulse response function. The results for GDP are largely the mirror image of those for unemployment. Relative to the no-shock baseline, real GDP starts to fall noticeably starting about two quarters after a contractionary shock. After 9 quarters, it is 4.4 percent below what it otherwise would have been. The maximum effect is again highly statistically significant ($t = -4.1$). The effects largely go away by five years after the shock.

One check that we think is valuable is to see if any one of our monetary shocks is driving the results. This is a way of gauging the importance of potential errors in the narrative classification—and more generally, of testing the robustness of the results. To do this, we just zero out the shocks one at a time and rerun the regressions. Figure 4 shows the results. There are ten lines, one for each variant of the shock dummy variable.

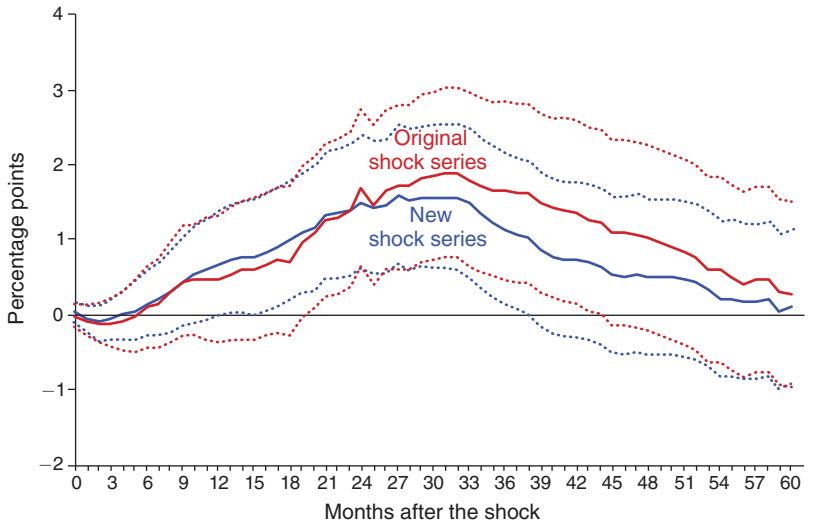

FIGURE 2. RESPONSE OF THE UNEMPLOYMENT RATE TO A MONETARY POLICY SHOCK, USING NEW AND ORIGINAL SHOCK SERIES

Notes: The figure shows the results of estimating equation (1) for horizons 0 to 60. The dependent variable is the unemployment rate. The dotted lines show the two-standard-error confidence bands. The new and original shock series are given in Table 2. See text for details of the estimation and the Data Appendix for the sources of the unemployment series.

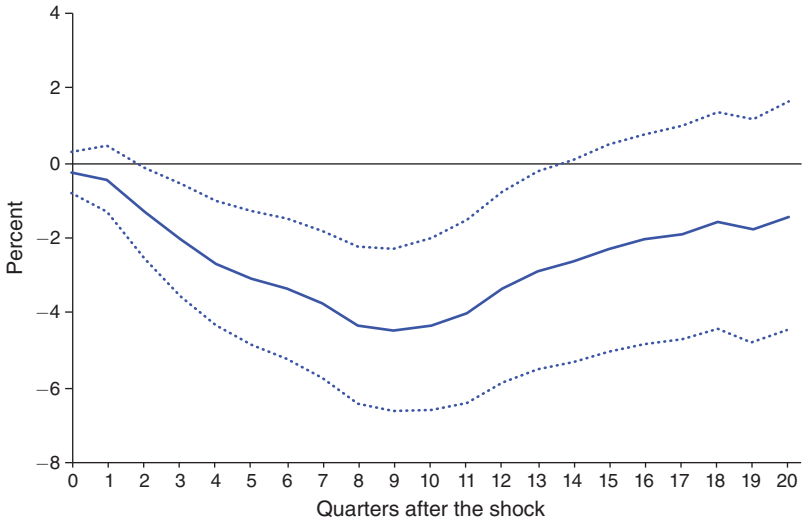

FIGURE 3. RESPONSE OF REAL GDP TO A MONETARY POLICY SHOCK

Notes: The figure shows the results of estimating equation (1) for horizons 0 to 20. The dependent variable is the log of real GDP. The dotted lines show the two-standard-error confidence bands. The new shock series is given in Table 2. See text for details of the estimation and the Data Appendix for the sources of the real GDP series.

The basic shape and magnitude of the effects are relatively unaffected by leaving out any one shock from our list. The negative impact of a shock is largest when we

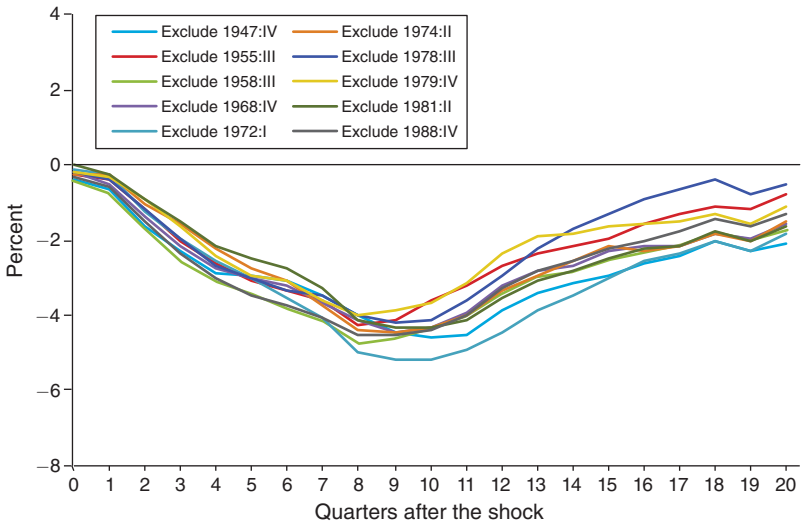

FIGURE 4. RESPONSE OF REAL GDP TO A MONETARY POLICY SHOCK, LEAVING OUT ONE SHOCK AT A TIME

Notes: The figure shows the results of estimating equation (1) for horizons 0 to 20 for ten variants of the new shock series. The variants are formed by sequentially eliminating one of the shocks. The dependent variable is the log of real GDP. The new shock series is given in Table 2. See text for details of the estimation and the Data Appendix for the sources of the real GDP series.

leave out the one positive shock—because the 1972 expansionary shock was followed not long after by the 1973 oil embargo and the related recession. It is smallest when we leave out the first Volcker shock. The impact at the five-year horizon goes away most completely when we leave out the 1978 shock. In all cases, the peak effect remains highly statistically significant. But the most striking thing is just how similar the ten lines are, particularly over the first eight quarters after the shock. This gives us confidence that the results are robust to possible errors in the identification of shocks. The same exercise for the monthly unemployment rate (presented in online Appendix B) also shows that the results are very similar across the ten variants of the shock series.

Inflation.—Something we didn't do in our original paper was to look at the behavior of inflation.[10] Our contractionary shocks are decisions to tighten policy because the current level of inflation was felt to be unacceptable. Our expansionary shock is a decision to loosen because the current stable rate of unemployment was felt to be unacceptable, realizing that such a policy risked raising inflation. It is natural to wonder what actually happened to inflation following these shocks. Did inflation fall after contractionary shocks and rise after expansionary ones?[11]

[10] Shapiro (1994) looks at the response of inflation to the original Romer and Romer dates. He finds that inflation falls following contractionary monetary shocks, but that the persistence and significance of the decline depends on the specification of the regression and the sample period.

[11] Importantly, our shocks are not responses to anticipated changes in inflation, but rather to the prevailing stable level. Thus, omitted variable bias is unlikely to be a substantial worry for the regression of inflation on our shock variable. Moreover, our regressions control for lags of inflation, and thus for the usual dynamics of inflation.

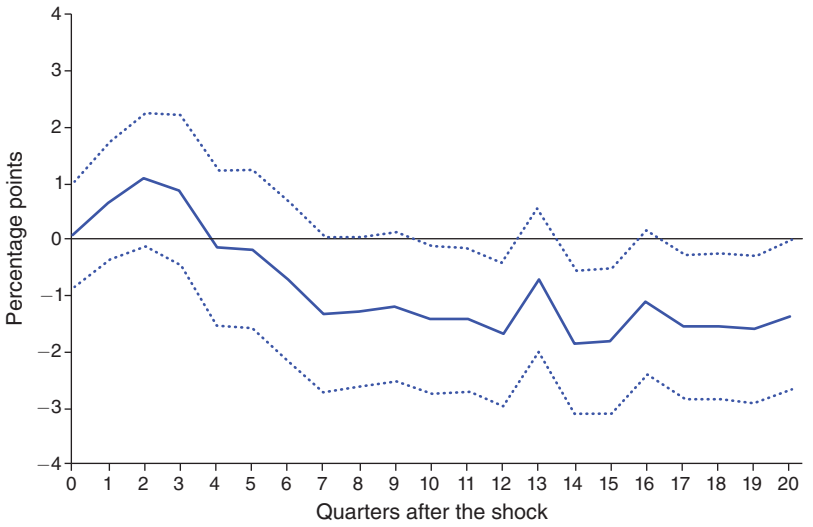

FIGURE 5. RESPONSE OF GDP PRICE INDEX INFLATION TO A MONETARY POLICY SHOCK

Notes: The figure shows the results of estimating equation (1) for horizons 0 to 20. The dependent variable is the inflation rate measured using the GDP price index. The dotted lines show the two-standard-error confidence bands. The new shock series is given in Table 2. See text for details of the estimation and the Data Appendix for the sources of the GDP price index series.

To answer this question, we run a similar empirical exercise, but with a measure of inflation on the left-hand side. For the baseline estimates, we measure inflation using the quarterly GDP price index. Specifically, we compute the change in the logarithm of the price index from the previous quarter times 400, so that it corresponds to inflation at an annual rate. The sequence of coefficients on the shock variable shows the impact of a monetary policy shock on inflation in percentage points, relative to the no-shock baseline.

Figure 5 shows the response of inflation to a contractionary monetary policy shock. There is evidence of a price puzzle at short horizons—that is, inflation initially rises somewhat (although the null hypothesis of no effect cannot be rejected). The point estimates turn negative four quarters after the shock—that is, inflation begins to fall below the baseline path one year after the shock. Inflation continues to fall over the second and third years after the shock, and then levels off. A contractionary monetary policy shock leads to a permanent reduction in inflation of about 1.5 percentage points. The estimated impact is significant at many horizons, but not wildly so. The largest t-statistic is 2.9, and the average t-statistic in years 3 through 5 is 2.2.

As with GDP, we consider the effect of leaving out one shock at a time on the estimated impulse response function for GDP price index inflation. To do this, we again construct ten variants of our shock series and estimate the sequence of regressions for each variant. Figure 6 shows the results. The ten estimated response functions share two common features. Each shows the inflation rate initially rising somewhat after the shock, and then falling relative to the baseline starting roughly one year after the shock. There is, however, decided variation across the ten variants—particularly in

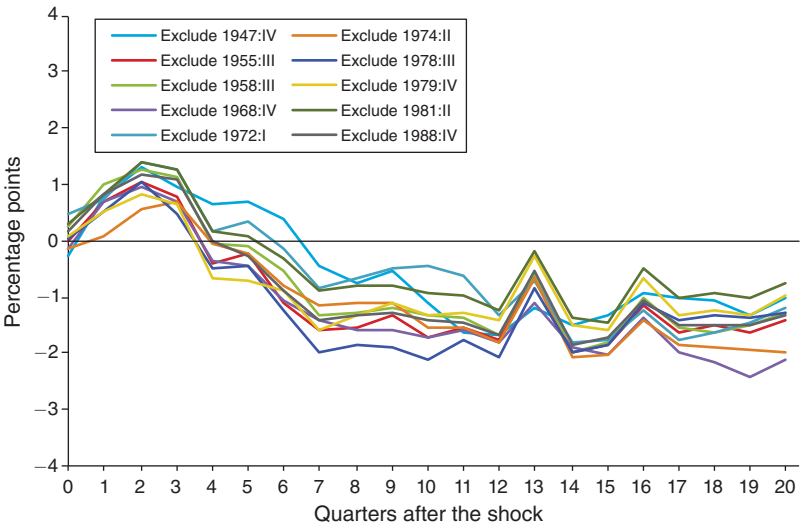

FIGURE 6. RESPONSE OF GDP PRICE INDEX INFLATION TO A MONETARY POLICY SHOCK, LEAVING OUT ONE SHOCK AT A TIME

Notes: The figure shows the results of estimating equation (1) for horizons 0 to 20 for ten variants of the new shock series. The variants are formed by sequentially eliminating one of the shocks. The dependent variable is the inflation rate measured using the GDP price index. The new shock series is given in Table 2. See text for details of the estimation and the Data Appendix for the sources of the GDP price index series.

the size of the inflation reduction in quarters 4 to 12 after the shock. This suggests that the inflation response varies noticeably across episodes.

We also consider the response of inflation measured using the PCE price index and the price index for PCE excluding food and energy (core PCE), both at a quarterly frequency. Figure 7 shows the results. The response of PCE inflation to a contractionary monetary policy shock (the blue line) is again slightly positive at short horizons (but not statistically significant), and then negative starting after one year. The peak effect is a reduction in inflation of 1.83 percentage points ($t = -2.4$) after 9 quarters. The results for core PCE inflation (the red line) are similar, but slightly weaker. The price puzzle is more pronounced at short horizons, and the negative impact at longer horizons is slightly smaller and somewhat more volatile.

Overall, the results for inflation suggest that Federal Reserve decisions to lower inflation do indeed reduce inflation. However, the results are not as definitive as those for real activity.

C. *Robustness*

In online Appendix B, we report the results of a large number of checks of our baseline specifications. Most importantly, we consider several additional outcome measures. For real activity, we consider monthly data on employment and industrial production. On the inflation side, we consider inflation measured using the monthly PCE price index and the core PCE price index. In addition, we consider multiple variations on our baseline specification and various ways of computing standard errors. Finally, as a simple way of visualizing what happened after the shocks (and

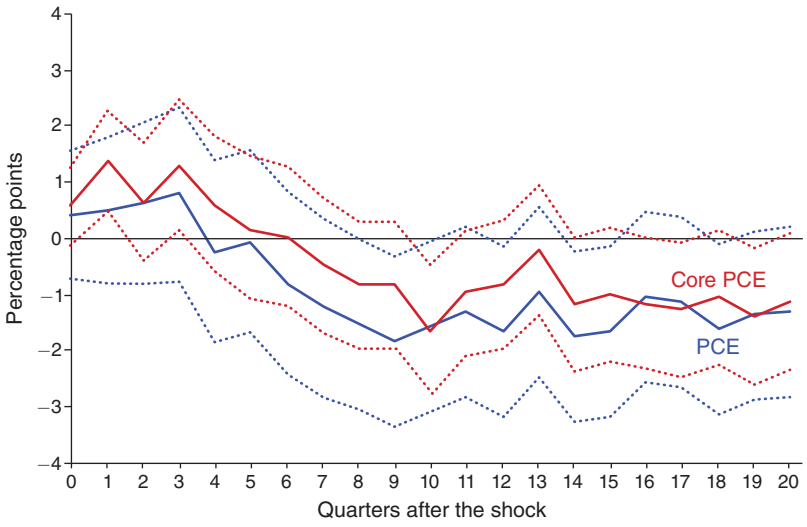

FIGURE 7. RESPONSE OF PCE AND CORE PCE INFLATION TO A MONETARY POLICY SHOCK

Notes: The figure shows the results of estimating equation (1) for horizons 0 to 20. The dependent variable is the inflation rate measured using the PCE price index and the core PCE price index. The dotted lines show the two-standard-error confidence bands. The new shock series is given in Table 2. See text for details of the estimation and the Data Appendix for the sources of the PCE and core PCE price index series.

analogous to some of the evidence we emphasize in Romer and Romer 1989), we examine forecast errors after each shock from simple univariate forecasting equations for real GDP, GDP price index inflation, and the unemployment rate.

The overarching message from these exercises is that the results are very robust. The specifics, of course, vary across the tests, but the changes are never large—and usually small—and are not systematic. For the additional outcome variables, the results for employment are very similar in terms of time pattern and statistical significance to those for the unemployment rate. The largest difference is that the impact of a monetary shock on employment fades considerably less completely than the impact on unemployment. The estimated response of industrial production looks very similar to that of real GDP. Finally, when we move from quarterly to monthly inflation data, the main changes are that the point estimates are more variable and the standard errors are somewhat larger. These are both natural consequences of the fact that monthly inflation is more volatile.

The numerous alternative specifications we consider yield only minor variation in the results. And the alternative approaches to computing standard errors do not yield results that differ systematically in either direction from our baseline (where we report conventional standard errors), and are sometimes irregular in ways that appear implausible.

The forecast errors for real GDP, while varying noticeably across episodes, consistently show a fall below (or, in the case of the positive shock in January 1972, a rise above) the path one would have expected based on a simple univariate forecast. The one exception is the September 1958 shock, where rapid growth in the quarter of the shock and the next three quarters pushed GDP well above the univariate

forecast path before an extended period of low and sometimes negative growth pushed it moderately below. The results for the unemployment rate are very similar to those for real GDP (with the opposite sign). For inflation, the results are much more mixed. For the most part, inflation moved in the expected direction—falling after the negative shocks and rising after the positive one. But there was a considerable lag after the August 1955 and October 1979 shocks. And after the shocks in December 1968, April 1974, and August 1978, inflation generally rose—often substantially. Again, this fits with the greater imprecision of our estimates of the impact of monetary policy on inflation.

III. Implications for Current Policy

It is widely understood that the Federal Reserve began tightening monetary policy in early 2022 because of concerns about inflation. What implications, if any, do our estimates of the impact of monetary policy shocks in the postwar period through 2016 have for current policy?

A. *Have We Had a Monetary Policy Shock?*

The obvious first question is whether the recent Federal Reserve actions constitute a monetary policy shock by our criteria. Importantly, we won't have the transcripts of the FOMC meetings for 2022 until 2028—so a definitive classification won't be possible for quite a while. But one thing we learned from our revisiting of the narrative evidence is that the "Record of Policy Actions" (very similar to what are now called the FOMC "Minutes") are a reasonably reliable source. And we do have those "Minutes" for the current period. We have read these following the approach and criteria used for identifying monetary shocks described in Section I.

Our tentative conclusion is that we have indeed had a contractionary monetary policy shock. American monetary policymakers have declared that the current level of inflation is unacceptable. They are taking actions to reduce it. And they are willing to accept substantial output and unemployment consequences to bring the reduction about. Here we describe some of the narrative evidence. A more detailed discussion of the evidence is contained in online Appendix A.

The FOMC began raising the federal funds rate in March 2022 because of its concern about inflation. But initially, monetary policymakers seemed to believe that they could get inflation down with few real effects. For example, the "Minutes" of the May meeting said, "members agreed that, with appropriate firming in the stance of monetary policy, they expected inflation to return to the Committee's 2 percent objective and the labor market to remain strong" ("Minutes," May 3–4, 2022, p. 9).[12] Because members did not express a willingness to accept output losses, the episode did not meet our requirements for a contractionary shock at this point.

The June FOMC meeting is the first time policymakers acknowledged possible negative real effects from the tightening to get inflation down. The "Minutes" said, "As the further firming in the policy stance would likely result in some slowing in

[12] All citations to the "Minutes" are to US Board of Governors of the Federal Reserve System (2022).

economic growth and tempering in labor market conditions, members also agreed to remove the previous statement language that had indicated an expectation that appropriate policy would result in a return of inflation to 2 percent and a strong labor market" ("Minutes," June 14–15, 2022, p. 10). In addition, "participants … anticipated that an appropriate firming of monetary policy would play a central role in helping address imbalances in the labor market. … [P]articipants generally expected the unemployment rate to increase" (p. 8). However, because policymakers did not go much beyond saying that the policy tightening would cause output to be lower than it otherwise would have been, we feel our criteria for monetary shock were not quite met at this meeting.

In July 2022, the FOMC more clearly met our criteria for a contractionary monetary policy shock. First, "Participants observed that inflation remained unacceptably high" ("Minutes," July 26–27, 2022, p. 8). This echoes one of our criteria for a contractionary shock almost word-for-word. Second, they were raising the funds rate to try to shift back the aggregate demand curve and induce a growth recession to reduce inflation. The "Minutes" reported, "the period ahead would likely see the response of aggregate demand to tighter financial conditions become stronger and more broadly based. Participants noted that a period of below-trend GDP growth would help reduce inflationary pressures and set the stage for the sustained achievement of the Committee's objectives of maximum employment and price stability" (p. 7). Third, FOMC members were willing to accept significant output costs to get inflation down: "Participants saw the risks to the outlook for real GDP growth as primarily being to the downside. These downside risks included the possibility that the tightening in financial conditions would have a larger negative effect on economic activity than anticipated" (p. 9). Because the statements are quite clear and emphatic, and confirm the vaguer statements from the June meeting, we think July is the most plausible date for the recent monetary policy shock.

The strongest statements consistent with there having been a contractionary shock come from the September FOMC meeting. FOMC members not only reiterated that inflation was unacceptably high and that they were willing to accept output losses to reduce it, they also expressed great resolve. The "Minutes" reported, "Participants reaffirmed their strong commitment to returning inflation to the Committee's 2 percent objective, with many stressing the importance of staying on this course even as the labor market slowed" ("Minutes," September 20–21, p. 9). Likewise, "A few participants … commented that the unemployment rate could rise by considerably more than in the staff forecast" (p. 8). This description echoed Federal Reserve Chair Jerome Powell's speech at Jackson Hole in late August, when he said, "While higher interest rates, slower growth, and softer labor market conditions will bring down inflation, they will also bring some pain to households and businesses. These are the unfortunate costs of reducing inflation" (Powell 2022, p. 1). It seems quite clear that a contractionary monetary shock occurred in the summer or early fall of 2022.

B. *Behavior of the Federal Funds Rate*

The behavior of the federal funds rate in the 2022 episode has been, if anything, a little more extreme than typical after a contractionary monetary policy shock.

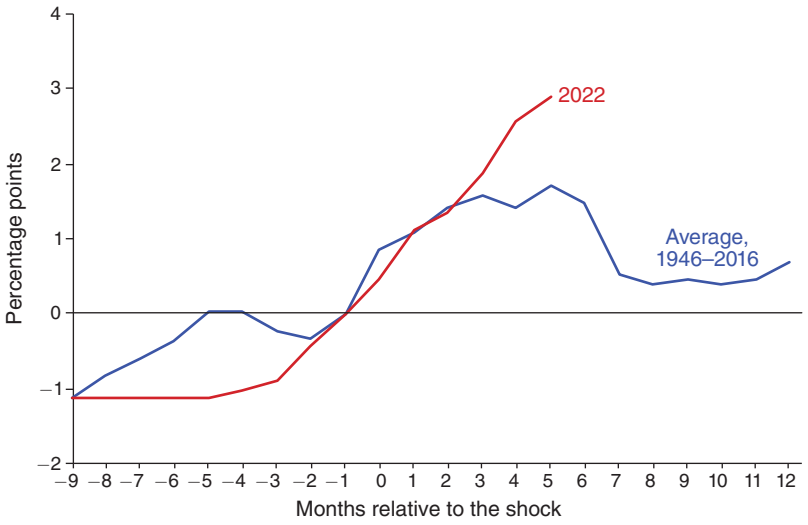

FIGURE 8. BEHAVIOR OF THE FEDERAL FUNDS RATE AROUND CONTRACTIONARY MONETARY POLICY SHOCKS AND IN 2022

Notes: The figure shows the average of the federal funds rate (relative to the month before the shock) for the nine contractionary monetary policy shocks between 1946 and 2016, as well as the federal funds rate (again relative to the month before the shock) for the shock tentatively dated in July 2022. The new dates of monetary policy shocks are given in Table 2. See text for details of the estimation and the Data Appendix for the sources of the federal funds rate series.

The blue line in Figure 8 shows the behavior of the funds rate, averaged over the nine contractionary shocks between 1946 and 2016. The funds rate is normalized to zero in the month before each shock. The funds rate typically rose some before the shocks. It has been common for policymakers to start to gradually tighten before they met all of our criteria for a contractionary shock. The funds rate then typically rose substantially after the shocks. On average, the funds rate rose a total of about 3 percentage points in a contractionary shock episode. The average obviously hides a great deal of variation across episodes—around the 1979 shock, the funds rate rose 7.5 percentage points; around the 1947 shock, our proxy for the funds rate (the three-month Treasury bill rate) rose less than 1 percentage point.

The red line in Figure 8 shows the behavior of the funds rate in the current episode. The funds rate in the current episode is normalized to zero in June 2022—one month before the candidate July 2022 date for the shock. As of early January 2023, the funds rate has risen just over 4 percentage points. About 1 percentage point of that occurred before the shock and 3 percentage points have occurred after. So the rise is a little larger than the average so far. It is roughly on par with the 1988 shock, the last shock in our sample.

C. *Possible Impacts*

Sitting in January 2023—two quarters after the date of the most recent contractionary monetary policy shock—what, if anything, does the experience following previous monetary shocks tell us about what to expect?

One is that we should not expect inflation to fall rapidly. The impulse response function for inflation shown in Figure 5 indicates that inflation would not normally have begun to drop two quarters after a contractionary shock. Now, we could get lucky. Positive supply shocks such as a fall in the price of oil could cause inflation to fall faster. We have seen some of this already. It could also be that inflation is more sensitive to expectations today than in the past, and aggressive Federal Reserve action could lower expected inflation. This, too, could cause inflation to fall more rapidly in the current episode than was typical in the past. More generally, there is substantial uncertainty about the effects of monetary shocks on inflation, so making predictions is difficult. But based on the historical point estimates, to get inflation down just through the conventional effects of contractionary monetary policy is likely to take at least another year.

The second thing that our analysis of past monetary shocks suggests is that as of January 2023, the effects on unemployment are likely yet to come. The impulse response function for the unemployment rate in Figure 1 shows monetary policy affects unemployment slowly. At six months after a shock, which is where January 2023 stands relative to the tentatively identified shock in July 2022, unemployment normally would not have begun to rise. If the historical pattern holds, the effects on unemployment would develop gradually over 2023.

The third thing our analysis suggests is that it is possible to go too far with monetary tightening. Because of the lags involved, policymakers will face a difficult decision about when to stop rate increases or reverse course. If policymakers keep tightening until inflation falls as much as they want, they will likely have gone too far—because the effects of tight policy will continue for many months after they stop raising rates. Now, how much higher rates need to go and how long they need to stay elevated is hard to say. In the modern era, it seems to take larger interest rate changes to get a given movement in the real economy. So, even if the federal funds rate has risen more than the historical average, monetary policymakers may need to raise it further. But policymakers are going to need to dial back monetary contraction before the inflation problem is completely solved, if they want to get inflation down without causing more pain than necessary.

IV. The Future of Narrative Research

So far, we have described how to make narrative research rigorous, and discussed the results of our application to estimating the effects of monetary policy—both for the past and the present. We conclude by considering the future of narrative research.

A. *Do We Still Need the Narrative Approach?*

A key goal of the narrative approach is to help to deal with omitted variable bias in empirical macroeconomic research. An obvious question is, haven't we come up with new and better ways of solving the identification problem in macroeconomics? Our answer is, not entirely.

One thing that macroeconomists have done is to appropriate the techniques of our micro colleagues. Using cross-section evidence to tease out macroeconomic relationships has been a huge step forward. For example, in a very nice paper, Gabriel

Chodorow-Reich and others (2012) use the relatively random variation in part of the Obama stimulus across states to estimate the causal impact of government spending on output and employment.

But this approach isn't a panacea. Cross-section evidence is inherently about the cross section. We need assumptions or aggregate evidence to understand how the cross-section evidence translates into the time-series relationship we actually care about.[13] Moreover, there are some issues where using cross-section evidence is especially challenging. The effect of monetary policy is one—monetary policy just doesn't vary at the micro level. So there will necessarily remain an important role for aggregate evidence, and thus room for the narrative approach in macroeconomics.

Another approach macroeconomists have come to rely on for identification is high-frequency evidence. If we have hourly or minute-by-minute data on expectations about the federal funds rate, we can use changes in a short window around monetary policy announcements to identify policy surprises. And, a number of studies use these surprises to try to estimate the effects of monetary policy. For example, in an early study using this approach, John Cochrane and Monica Piazzesi (2002) find a statistically insignificant rise in employment following a contractionary monetary policy shock.

This approach also has considerable value, especially for studying the effects of monetary policy on financial variables. But it too is not a panacea. One problem is that such surprises tend to be very small. Modern monetary policymakers typically signal changes in policy well in advance, so we are trying to measure effects off quite tiny blips. For example, Emi Nakamura and Jón Steinsson (2018) report that the standard deviation of the high-frequency monetary shocks they find around announcements following FOMC meetings is only five basis points. More fundamentally, the response to these monetary policy surprises may not capture the true causal impact of monetary policy. It may be that well-signaled interest rate movements have larger or smaller real effects than surprise ones. Or, it could be that what the market views as a surprise is actually monetary policymakers' response to something else that also affects the real economy that they know about but others do not. Thus, the response to such a surprise could capture the impact of the other factor, not monetary policy.[14] So once again, using narrative evidence on the motivation of policymakers may be a very useful alternative for capturing the true causal impact of monetary policy decisions.

Third, macroeconomists have come up with various purely econometric approaches that try to achieve identification, such as simple vector autoregressions (VARs). At its most basic level, a VAR posits a system of variables—say interest rates, output, and prices. Each variable is assumed to depend on lags of itself, and on the contemporaneous and lagged value of other variables. The identifying assumptions typically involve timing. For example, in the case of estimating the effects of monetary policy, interest rates are typically assumed to respond to output contemporaneously, but output is assumed to only respond to interest rates with a lag.

Unfortunately, this approach doesn't deal with forward-looking policy. If the central bank responds to information beyond the variables in the VAR about the

[13] This is a point made very persuasively by Nakamura and Steinsson (2014).
[14] See Romer and Romer (2000) and Nakamura and Steinsson (2018).

future path of output, there will still be omitted variable bias. And central banks typically invest lots of resources in forecasting the economy, precisely so that they can move interest rates in anticipation of future movements in output and inflation. Forward-looking policy is not a problem that can be solved by simple or even sophisticated timing assumptions. As a result, all of the purely econometric advances still need something else to deal with omitted variable bias.

Finally, it is important to realize that some macroeconomic behavior may be fundamentally episodic in nature. Financial crises, recessions, disinflations, are all events that seem to play out in an identifiable pattern. There may be long periods where things are basically fine, that are then interrupted by short periods when they are not. If this is true, the best way to understand them may be to focus on episodes—not a cross-section proxy or a tiny sub-period. In addition, it is valuable to know when the episodes were and what happened during them. And, the identification and understanding of episodes may require using sources other than conventional data.

For all of these reasons, we think there are a number of questions in macroeconomics where using the plethora of qualitative data available is still valuable.

B. *Can Narrative Research Be Delegated to Research Assistants or Computers?*

Even if the narrative approach is still useful, does it need to be done the way we do it? That is, does it need to be done by scholars reading the documents themselves? That way is certainly time consuming (one might even go so far as to say tedious). Could the narrative analysis be delegated to research assistants (RAs) or computers? It is useful to consider the two alternatives together because many of the same issues arise in both cases. Our answer is a qualified yes—at some point, in some cases, if it is done very carefully.

First, a key issue is how sophisticated is the information one is trying to get from the narrative source? Narrative evidence takes many forms. It may be a fairly straightforward series on the severity of financial crises deduced from the reports of the OECD or the IMF (see Romer and Romer 2017). It may be a more subtle series on when there was news about defense spending pieced together from many news sources and government documents (see Ramey 2011). Or it may be a complicated classification of when monetary policy is likely to be uncorrelated with other factors driving real economic activity, as in this study.

As a general matter, simpler tasks are easier to delegate to RAs or computers than more complex questions. For example, classifying whether the tone of the discussion at a Federal Reserve meeting was positive or negative is surely much easier to delegate than whether monetary policy decisions were a response to the state of the economy.

But even in straightforward tasks, computers and untrained research assistants may leave much to be desired. In a study of exchange rate intervention, Alain Naef (2022) considers three approaches to classifying contemporaneous daily narrative descriptions of the forces affecting the exchange rate into positive, neutral, and negative: his own (presumably expert) reading, the judgment of participants on Amazon Mechanical Turk (an on-demand, task-based labor website), and the results of a natural language processing algorithm. Although the results from both Amazon Mechanical Turk and the natural language processing are positively correlated with Naef's classification, the correlations are only modest.

Similarly, Adam Shapiro and Daniel Wilson (2022) use a computer algorithm to classify the negativity of comments by Federal Reserve policymakers. How concerned do policymakers seem to be about the economic outlook? They compare the results of the algorithm to trained human auditors. They find that the rank correlation coefficient is only about 0.4. Clearly, the two methods classify the texts quite differently.

A crucial step in successfully delegating the narrative analysis to either human RAs or computers is training. As we discussed earlier, figuring out what one is trying to get out of a narrative source is a crucial and difficult step. It is hard to imagine a computer ever providing that creative spark. What is more likely to work is for researchers to figure out what can be learned from the source and how to classify it, and then train others—human or computer.

Such training will surely involve the researchers doing some narrative analysis and classification themselves, and then teaching others to match their procedures and results. For example, it would be interesting to know whether looking at where the human readers and the computer algorithms used by Shapiro and Wilson and by Naef went astray, and providing training to avoid those pitfalls, would lead to greatly improved outcomes. Now, extensive training has the drawback that it risks replicating some of the biases of the original researchers. But for complex narrative analysis, it is likely to be essential.

Even then, machine learning is unlikely to work in some situations. One is when there are a limited number of observations from which the computer can learn. It is hard to imagine that we could train a computer to read Federal Reserve transcripts the way we do, when we only have ten examples of monetary shocks in 70 years. Another time when machine learning is unlikely to work is when the categorization of the narrative source is highly complex or subtle. Overall, there are certainly ways to delegate much of the narrative work and thus make it easier and more reproducible. But, we suspect that full mechanization still has a long way to go.

For now, the main way to use the narrative approach is going to be the old-fashioned one—scholars sitting at their desks (or in our case, our dining room table) doing a lot of careful reading of the narrative sources themselves. Or, to put it in more personal terms, we thoroughly expect to be made largely redundant by computers eventually, but perhaps not for a few years to come.

Data Appendix

The data series we use are the conventional government statistics for key macroeconomic indicators. All series are seasonally adjusted. Because we want to use data back to 1946 in our estimation, we often need to construct a proxy for the official series for the very beginning of the sample. This Appendix describes our data sources and explains our procedures for constructing those proxies.

A. *Unemployment Rate*

The monthly unemployment rate data for 1948:1 to 2016:12 are from the US Bureau of Labor Statistics (BLS), series LNS14000000, seasonally adjusted, percent, downloaded November 20, 2022. All BLS data are from https://www.bls.gov/data.

To construct a proxy for the standard series for 1946:10 to 1947:12, we use data from the Current Population Reports produced by the Bureau of the Census (Census). The data are conceptually similar to the data collected starting in 1948 by the BLS through the Current Population Survey. The data for 1946:10–1946:12 are from US Bureau of the Census (1947, p. 17); the data for 1947 and 1948 (data for 1948 are used for comparison with the BLS series) are from the US Bureau of the Census (1948, p. 174 and 1949, p. 174). The unemployment rate is calculated as the ratio of the number unemployed to the total civilian labor force. Because the labor force and unemployment data are not seasonally adjusted, this calculated unemployment rate is unadjusted as well.

To seasonally adjust the Census series for 1946:10–1948:12, we use the difference between the BLS seasonally adjusted unemployment rate series and the unadjusted BLS series in each month of 1948 to create monthly seasonal adjustment factors. The unadjusted unemployment rate data for 1948:1 to 1948:12 are from the BLS, series LNU04000000, not seasonally adjusted, percent, downloaded January 23, 2023.

We add these monthly adjustment factors to the Census series for 1946:10–1948:12 to create a seasonally adjusted series. We then calculate the average difference between the seasonally adjusted BLS series and our constructed seasonally adjusted Census series in 1948. We then add this difference (which is 0.4 percentage point) to the observations before 1948.

B. *Real GDP*

The quarterly real GDP data for 1947:I to 2016:IV are from the US Bureau of Economic Analysis (BEA), Table 1.1.6, series gross domestic product, seasonally adjusted, billions of chained (2012) dollars, annual rates, downloaded November 20, 2022. All BEA data are from https://www.bea.gov/itable/national-gdp-and-personal-income.

To construct a proxy for real GDP for 1946:IV, we simply take the ratio of real GDP to the Index of Industrial Production (at a quarterly frequency) in 1947:I, and multiply it by the Index of Industrial Production in 1946:IV. That is, we assume that the percentage change in real GDP between 1946:IV and 1947:I is the same as the percentage change in industrial production. The data for industrial production are from the US Board of Governors of the Federal Reserve System, G.17 Industrial Production and Capacity Utilization, series IP.B50001.S, seasonally adjusted, index, 2017=100, downloaded November 20, 2022. We convert the monthly index to quarterly by averaging. All Board of Governors data are from https://www.federalreserve.gov/datadownload.

C. *GDP Price Index*

The quarterly GDP price index data for 1947:I to 2016:IV are from the BEA, Table 1.1.4, series gross domestic product, seasonally adjusted, index, 2012=100, downloaded November 20, 2022.

Because we specify the variable in percentage changes in the estimation, we need to construct a proxy for the level of the index for both 1946:III and 1946:IV. To do

this, we simply take the ratio of the GDP price index to the Consumer Price Index for All Urban Consumers (CPI-U) at a quarterly frequency in 1947:I, and multiply it by the quarterly CPI-U in 1946:III and 1946:IV. The CPI-U data are from the BLS, series CUUR0000SA0, not seasonally adjusted, index, 1982–84=100, downloaded November 20, 2022. We convert the monthly series to quarterly by averaging. The CPI-U for 1946 is only available in seasonally unadjusted form. However, because seasonal movements in the CPI-U are relatively minor, we make no further adjustment to our constructed proxy. Inflation at an annual rate is calculated as the difference in logarithms times 400.

D. *PCE Price Index and PCE Price Index Excluding Food and Energy*

The quarterly data on the personal consumption expenditures (PCE) price index for 1947:I to 2016:IV and the PCE price index excluding food and energy for 1959:I to 2016:IV are from the BEA, Table 2.3.4, series personal consumption expenditures and PCE excluding food and energy, seasonally adjusted, index, 2012=100, downloaded November 20, 2022.

To construct a proxy for the PCE price index for 1946:III and 1946:IV, we take the ratio of the PCE price index to the CPI-U at a quarterly frequency in 1947:I, and multiply it by the quarterly CPI-U in 1946:III and 1946:IV (which, as described above, is only available seasonally unadjusted). The CPI-U data are from the BLS, series CUUR0000SA0, not seasonally adjusted, index, 1982–84=100, downloaded November 20, 2022. We convert the monthly series to quarterly by averaging.

To construct a proxy for the PCE price index excluding food and energy for 1957:I to 1958:IV, we take the ratio of the PCE price index excluding food and energy to the CPI-U less food and energy at a quarterly frequency in 1959:I, and multiply it by the quarterly CPI-U less food and energy. The CPI-U less food and energy data are from the BLS, series CUSR0000SA0L1E, seasonally adjusted, index, 1982–84=100, downloaded November 20, 2022. To construct a proxy for the PCE price index excluding food and energy for 1947:I to 1956:IV, we take the ratio of our proxy to the quarterly CPI-U less food in 1957:I, and multiply it by the CPI-U less food. The CPI-U less food data are from the BLS, series CUSR0000SA0L1, seasonally adjusted, index, 1982–84=100, downloaded November 20, 2022. Finally, to construct a proxy for the PCE price index excluding food and energy for 1946:III and 1946:IV, we take the ratio of our proxy to the quarterly CPI-U less food (which is only available seasonally unadjusted for this period) in 1947:I, and multiply by the CPI-U less food. The CPI-U less food data are from the BLS, series CUUR0000SA0L1, not seasonally adjusted, index, 1982–84=100, downloaded November 20, 2022. We convert the various monthly CPI series to quarterly by averaging.

Inflation at an annual rate is calculated as the difference in logarithms times 400.

E. *Federal Funds Rate*

The monthly effective federal funds rate data for 1954:7 to 2022:12 are from the US Board of Governors of the Federal Reserve System, H.15 Selected Interest Rates, series RIFSPFF_N.M, percent per year, downloaded January 15, 2023.

To construct a proxy for the effective federal funds rate for 1946:10 to 1954:6, we take the difference between the effective federal funds rate and the three-month Treasury bill rate in 1954:7 (which is 0.08 percentage point), and then add that to the three-month Treasury bill rate. The monthly data for the three-month Treasury bill rate are from the US Board of Governors of the Federal Reserve System, H.15 Selected Interest Rates, series RIFSGFSM03_N.M, three-month Treasury bill secondary market rate discount basis, percent per year, downloaded November 27, 2022.

REFERENCES

Chodorow-Reich, Gabriel, Laura Feiveson, Zachary Liscow, and William Gui Woolston. 2012. "Does State Fiscal Relief during Recessions Increase Employment? Evidence from the American Recovery and Reinvestment Act." *American Economic Journal: Economic Policy* 4 (3): 118–45.

Cochrane, John H., and Monika Piazzesi. 2002. "The Fed and Interest Rates: A High-Frequency Identification." *American Economic Review* 92 (2): 90–95.

Friedman, Milton, and Anna Jacobson Schwartz. 1963. *A Monetary History of the United States: 1863–1960*. Princeton: Princeton University Press.

Goodfriend, Marvin, and Robert G. King. 2005. "The Incredible Volcker Disinflation." *Journal of Monetary Economics* 52 (5): 981–1015.

Hamilton, James D. 1983. "Oil and the Macroeconomy since World War II." *Journal of Political Economy* 91 (2): 228–48.

Hamilton, James D. 1985. "Historical Causes of Postwar Oil Shocks and Recessions." *Energy Journal* 6 (1): 97–116.

Hansen, Stephen, Michael McMahon, and Andrea Prat. 2018. "Transparency and Deliberation within the FOMC: A Computational Linguistics Approach." *Quarterly Journal of Economics* 133 (2): 801–70.

Hetzel, Robert L. 2008. *The Monetary Policy of the Federal Reserve*. Cambridge: Cambridge University Press.

Jalil, Andrew J. 2015. "A New History of Banking Panics in the United States, 1825–1929: Construction and Implications." *American Economic Journal: Macroeconomics* 7 (3): 295–330.

Jordà, Òscar. 2005. "Estimation and Inference of Impulse Responses by Local Projections." *American Economic Review* 95 (1): 161–82.

Kareken, John, and Robert M. Solow. 1963. "Lags in Monetary Policy." In *Stabilization Policies,* edited by E. Cary Brown, 14–96. Englewood Cliffs, NJ: Prentice-Hall.

Meade, Ellen E., and David Stasavage. 2008. "Publicity of Debate and the Incentive to Dissent: Evidence from the US Federal Reserve." *Economic Journal* 118 (528): 695–717.

Meltzer, Allan H. 2003. A *History of the Federal Reserve: Volume I, 1913–1951*. Chicago: University of Chicago Press.

Meltzer, Allan H. 2009. *A History of the Federal Reserve: Volume II, Book 2, 1970–1986.* Chicago: University of Chicago Press.

Naef, Alain. 2022. "Blowing against the Wind? A Narrative Approach to Central Bank Foreign Exchange Intervention." Unpublished.

Nakamura, Emi, and Jón Steinsson. 2014. "Fiscal Stimulus in a Monetary Union: Evidence from US Regions." *American Economic Review* 104 (3): 753–92.

Nakamura, Emi, and Jón Steinsson. 2018. "High-Frequency Identification of Monetary Non-Neutrality: The Information Effect." *Quarterly Journal of Economics* 133 (3): 1283–330.

Nelson, Edward. 2005. "The Great Inflation of the Seventies: What Really Happened?" *B.E. Journal of Macroeconomics: Advances in Macroeconomics* 5 (1): 1–48.

Powell, Jerome H. 2022. "Monetary Policy and Price Stability." Speech, Federal Reserve Bank of Kansas City Economic Policy Symposium, Jackson Hole, WY, August 26, 2022. https://www.federalreserve.gov/newsevents/speech/powell20220826a.htm.

Ramey, Valerie A. 2011. "Identifying Government Spending Shocks: It's All in the Timing." *Quarterly Journal of Economics* 126 (1): 1–50.

Romer, Christina D., and David H. Romer. 1989. "Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz." *NBER Macroeconomics Annual* 4: 121–70.

Romer, Christina D., and David H. Romer. 1994. "Monetary Policy Matters." *Journal of Monetary Economics* 34 (1): 75–88.

Romer, Christina D., and David H. Romer. 2000. "Federal Reserve Information and the Behavior of Interest Rates." *American Economic Review* 90 (3): 429–57.

Romer, Christina D., and David H. Romer. 2010. "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks." *American Economic Review* 100 (3): 763–801.

Romer, Christina D., and David H. Romer. 2017. "New Evidence on the Aftermath of Financial Crises in Advanced Economies." *American Economic Review* 107 (10): 3072–118.

Romer, Christina D., and David H. Romer. 2023. "Replication Data for Presidential Address: Does Monetary Policy Matter? The Narrative Approach after 35 Years." American Economic Association [Publisher], Inter-university Consortium for Political and Social Research [Distributor]. https://doi.org/10.3886/E185843V1.

Shapiro, Adam Hale, and Daniel J. Wilson. 2022. "Taking the Fed at its Word: A New Approach to Estimating Central Bank Objectives Using Text Analysis." *Review of Economic Studies* 89 (5): 2768–805.

Shapiro, Matthew D. 1994. "Federal Reserve Policy: Cause and Effect." In *Monetary Policy*, edited by N. Gregory Mankiw, 307–32. Chicago: University of Chicago Press.

US Board of Governors of the Federal Reserve System. 1946–1992. "Record of Policy Actions of the Federal Open Market Committee."

US Board of Governors of the Federal Reserve System. 1946–1967. *Minutes of Federal Open Market Committee.*

US Board of Governors of the Federal Reserve System. 1967–1976. *Memoranda of Discussion of the Federal Open Market Committee*.

US Board of Governors of the Federal Reserve System. 1976–2016. *Transcripts of the Federal Open Market Committee.*

US Board of Governors of the Federal Reserve System. 2022. "Minutes of the Federal Open Market Committee."

US Board of Governors of the Federal Reserve System. 2022–2023. "Data Download Program." https://www.federalreserve.gov/datadownload.

US Bureau of Economic Analysis. 2022. "National GDP & Personal Income." https://www.bea.gov/itable/national-gdp-and-personal-income.

US Bureau of Labor Statistics. 2022–2023. "Databases, Tables & Calculators by Subject." https://www.bls.gov/data.

US Bureau of the Census. 1947. "Labor Force, Employment, and Unemployment in the United States, 1940–1946." *Current Population Reports: Labor Force Bulletin.* Series P-50, No. 2.

US Bureau of the Census. 1948–1949. *Statistical Abstract of the United States.* Washington, DC: US Government Printing Office.