NBER WORKING PAPER SERIES

THE MACROECONOMIC EFFECTS OF TARIFFS:
INSIGHTS FROM 180 YEARS OF U.S. TRADE POLICY

Tamar den Besten
Regis Barnichon
Diego R. Känzig
Aayush Singh

Working Paper 35102
http://www.nber.org/papers/w35102

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2026

This paper unifies and extends two prior working papers: “The Macroeconomic Effects of Tariffs: Evidence from U.S. Historical Data” (Den Besten and Känzig) and “What Is a Tariff Shock? Insights from 150 Years of Tariff Policy” (Barnichon and Singh). We thank Marios Angeletos, Javier Bianchi, Larry Christiano, Marty Eichenbaum, Andrew Fieldhouse, Luca Fornaro, Yuriy Gorodnichenko, Lukas Hack, Douglas Irwin, Joel Mokyr, Emi Nakamura, Giorgio Primiceri, Valerie Ramey, Matt Rognlie, Kunal Sangani, Jón Steinsson, Ludwig Straub and Chris Wolf for helpful comments and discussions. The views expressed in this paper are the sole responsibility of the authors and do not necessarily reflect the views of the Federal Reserve Bank of San Francisco, the Federal Reserve System, or the National Bureau of Economic Research.

The Macroeconomic Effects of Tariffs: Insights from 180 Years of U.S. Trade Policy
Tamar den Besten, Regis Barnichon, Diego R. Känzig, and Aayush Singh
NBER Working Paper No. 35102
April 2026
JEL No. E30, F13, F14, F41, H20

ABSTRACT

We study the macroeconomic effects of tariff policy using U.S. historical data from 1840–2024. We construct a narrative series of plausibly exogenous tariff changes – based on major legislative actions, multilateral negotiations, and temporary surcharges – and use it as an instrument to identify a structural tariff shock. Tariff increases are contractionary: imports fall sharply, exports decline with a lag, and output and manufacturing activity drop persistently. The shock transmits through both supply and demand channels. Prices rise in the full sample but fall post-World War II, a pattern consistent with changes in the monetary policy response and with stronger international retaliation and reciprocity in the modern trade regime.

Tamar den Besten
Northwestern University
tamardenbesten2029@u.northwestern.edu

Regis Barnichon
Federal Reserve Bank of San Francisco
regis.barnichon@sf.frb.org

Diego R. Känzig
Northwestern University
Department of Economics
and NBER
dkaenzig@northwestern.edu

Aayush Singh
Federal Reserve Bank of San Francisco
aayush.singh@sf.frb.org

A narrative companion paper is available at http://www.nber.org/data-appendix/w35102

1. Introduction

The sharp increase in U.S. tariff rates in 2025 has renewed interest in a long-standing question: what are the macroeconomic effects of tariffs? Yet despite the renewed prominence of tariff policy, there is surprisingly little empirical evidence on its broader macroeconomic consequences. Most existing studies rely on post-World War II data – a period with relatively modest movements in average tariff rates – and focus primarily on partial-equilibrium or sectoral outcomes, such as relative price changes, industry-level effects, or bilateral trade flows. As a result, the aggregate effects of tariffs on output, prices, and trade remain less well understood.

While the current tariff increases are unprecedented relative to recent decades, historical experience offers episodes that resemble today's policy environment. In the pre-World War II era, U.S. tariff rates fluctuated sharply and frequently, driven by partisan and ideological conflict, fiscal pressures, and evolving international conditions. These forces generated large and rapid movements in trade barriers that resemble many features of contemporary tariff actions. In this paper, we exploit this rich historical variation to study the aggregate effects of tariff changes and to understand how tariff shocks propagate through the macroeconomy.

Identifying the causal effects of tariff changes from time-series variation requires isolating tariff movements that are not systematically shaped by macroeconomic conditions. Two major challenges arise. First, commonly used measures such as average tariff rates embed behavioral responses affecting imports and prices, so that changes in these measures may reflect economic developments rather than policy actions. This complicates the measurement of true policy-induced variation. Second, and more importantly, many historical tariff reforms were motivated by factors endogenous to the economy – including efforts to finance government spending or to respond to fiscal stress and economic downturns – making these tariff changes inappropriate for identifying exogenous shocks.

To overcome these challenges, we conduct a detailed narrative analysis of major U.S. tariff changes back to 1840. Drawing on authoritative historical accounts – most notably Irwin's (2017a) *Clashing over Commerce: A History of U.S. Trade Policy* – and complementing them with congressional records and the United States Statutes at Large, we identify 35 major tariff events over this period. These episodes include comprehensive congressional revisions of the tariff code, major multilateral tariff reductions, and broad-based temporary measures. Each event is documented with its legislative context, stated motivations, and implementation details, providing a rich foundation for distinguishing policy-driven actions from those tied to evolving macroeconomic conditions.

In a second step, we classify each event according to its underlying motivation, following the narrative approach (Romer and Romer, 2023). We label tariff changes as plausibly exogenous when they were driven by ideological commitments to protectionism or free trade, distributional motivations, or other factors unrelated to current or prospective real activity. By contrast, we classify tariff changes as endogenous when they were enacted to finance contemporaneous government spending, stabilize economic activity, or otherwise respond directly to prevailing macroeconomic conditions. This narrative classification isolates the subset of tariff actions suitable for identifying the causal effects of tariffs. Out of the 35 events we consider, we classify 21 as plausibly exogenous.

Building on this classification, we construct a narrative tariff shock series based on the date and sign of each plausibly exogenous event. This approach requires only that the direction and timing of these policy actions be plausibly exogenous, without imposing assumptions on the exogeneity of their size. Measuring the magnitude of tariff changes is challenging, as average tariff rates also reflect endogenous shifts in import composition and prices in addition to statutory changes. Focusing on the sign of each event therefore provides a cleaner measure of tariff policy shocks. We use this narrative shock series as an instrument in a semi-structural model of the U.S. economy to estimate the dynamic causal effects of tariff shocks.

Our results indicate that tariff shocks are contractionary for the U.S. economy. A tariff increase that raises the average tariff rate by 1 percentage point leads to a peak decline in real GDP of roughly 0.9 percent, with output remaining below trend for several years. Trade activity contracts markedly. Imports fall immediately by around 4 percent, while exports tend to increase slightly on impact but then also fall with a lag, reaching a maximum decline of about 2 percent. Manufacturing production declines even more strongly – by over 1.5 percent at peak – suggesting that tariff increases do not shield domestic industry in the aggregate, despite their protective intent. Consistent with the substantial drop in imports, the U.S. dollar appreciates. The price response is more nuanced: the GDP deflator increases by around 0.5 percent at peak, but the estimate is imprecise, and the effect gradually dissipates as the economy contracts.

Taken together, these results point to the presence of both supply and demand channels in the transmission of tariff shocks. In the full historical sample, the short-run increase in prices combined with falling output resembles a cost-push pattern consistent with higher input costs and reduced foreign competition. However, the persistent declines in output and trade volumes suggest that demand effects are also important, particularly through weaker exports and broader macroeconomic contraction. Thus, the simultaneous presence of supply-side inflationary pressures and demand-driven disinflation could explain the muted overall response of aggregate prices.

The behavior of the economy in different subsamples further clarifies these mechanisms. When we restrict the sample to end in 1945, the responses closely mirror the baseline, indicating that the stagflationary effects are not driven by postwar dynamics. By contrast, estimates based solely on the post-World War II period yield qualitatively different patterns: exports fall sharply on impact and prices decline following tariff shocks.

These differences point to stronger demand effects in the modern era and align with the shift away from the gold standard, which altered the exchange rate adjustment process and the transmission of trade policy to domestic prices. Under floating exchange rates, tariffs induce an immediate dollar appreciation, which partly offsets the tariff-driven increase in import prices, while simultaneously reducing export demand. Both these mechanisms exert downward pressure on domestic prices, relative to the gold standard period, and help explain the disinflationary response observed in the postwar sample. Additionally, tariff policy during much of the modern era has been coordinated through international negotiations, so changes frequently occurred in a multilateral context that may have amplified the demand channel through stronger foreign market responses. Indeed, we show that foreign tariff rates increase significantly following a U.S. tariff increase in the postwar sample, whereas the foreign response is more modest and statistically insignificant in the earlier period. Consistent with these mechanisms, we find a more pronounced output response in the postwar sample.

Accounting for the endogeneity of tariff policy is central to uncovering these effects. A simple recursive identification that treats all tariff movements as exogenous innovations yields markedly different responses. Similarly, estimates based on tariff reforms we classify as endogenous imply very different responses of output, prices, and trade flows, which do not resemble the effects of a tariff shock. This contrast underscores that tariff changes cannot be interpreted as structural shocks without carefully isolating the motivations behind them; the narrative approach is therefore essential for obtaining credible causal estimates.

We subject our findings to an extensive set of robustness and sensitivity checks. We assess the role of identifying assumptions, including tests for invertibility and alternative dynamic models; we examine the influence of confounding shocks and influential historical events, such as wartime and interwar tariff acts; and we account for anticipation by excluding episodes with long lags between signing and implementation. Across all exercises, the contractionary impact of tariff shocks remains strong and stable.

Overall, the central message is clear: tariff increases depress economic activity and trade once their indirect, general-equilibrium effects are taken into account. While tariffs may appear protective in partial equilibrium, they lead to sizable declines in output, manufacturing activity, and external competitiveness at the aggregate level. Our findings suggest that these effects transmit through both supply and demand channels, with their relative importance varying over time. In the postwar era in particular, changes in the monetary policy reaction and the reciprocal structure of the trading system strengthen demand-side effects and appear to have reshaped the transmission of tariff shocks toward weaker exports and disinflationary price responses.

Related literature and contribution. Our paper contributes to several strands of literature on trade policy and macroeconomic dynamics. A first strand studies the partial-equilibrium effects of recent U.S. tariff actions, most notably the 2018–2019 Trump tariffs. Amiti, Redding, and Weinstein (2019), Fajgelbaum et al. (2020), Flaaen and Pierce (2024), and Cox (2025) document substantial increases in import prices, reductions in trade volumes, and distributional consequences across firms and regions. These studies provide valuable evidence on the microeconomic and relative-price effects of tariffs, but they do not speak directly to the aggregate, general-equilibrium effects on output, inflation, and overall macroeconomic performance.

A second strand studies the macroeconomic consequences of tariff changes, exploiting time-series variation in the post-World War II period (Furceri et al., 2022; Boer and Rieth, 2024; Schmitt-Grohé and Uribe, 2025; Bandyopadhyay, Ferraro, and Bower, 2025; Franconi and Hack, 2025). These studies provide important evidence on how tariffs shape output, prices, and trade flows, but they face two potential limitations. First, postwar U.S. tariff variation is relatively modest and reflects mostly negotiated reductions rather than large unilateral changes. Second, tariff increases were particularly rare in this period, and when they did occur they were typically implemented in response to adverse economic conditions, such as balance-of-payments pressures in the case of the Nixon surcharge (Irwin, 2013), which complicates causal identification. Our narrative approach addresses these issues by isolating plausibly exogenous tariff actions over a much longer sample.

By leveraging historical variation, our paper relates to a growing body of macroeconomic work that uses long-run data to identify the effects of interest. Examples include Ramey and Zubairy (2018) and Antolin-Diaz and Surico (2025) on government spending, Jordà, Singh, and Taylor (2024) on monetary policy or Jordà, Schularick, and Taylor (2013, 2016) on the effects of credit booms and financial cycles. Our analysis follows this tradition by exploiting 180 years of U.S. tariff policy, where substantial prewar swings offer precisely the type of plausibly exogenous variation needed to credibly estimate the macroeconomic effects of tariff shocks. In this sense, we also contribute to a burgeoning economic history literature on tariffs (see e.g., O'Rourke, 2000; Clemens and Williamson, 2004; Schularick and Solomou, 2011; Klein and Meissner, 2024; Greenland and Lopresti, 2024).

From a methodological viewpoint, our paper builds on an influential literature that uses narrative information to isolate policy shocks. Seminal contributions by Romer and Romer (1989, 2004) demonstrate how narrative evidence can be used to separate exogenous policy actions from those responding to economic conditions in the context of monetary policy. This approach has since been extended to other domains including tax policy (Romer and Romer, 2010; Mertens and Ravn, 2013; Cloyne, 2013) and government spending (Ramey and Shapiro, 1998; Ramey, 2011; Fieldhouse and Mertens, 2023). We port this narrative methodology to tariff policy, constructing a new long-run series of plausibly exogenous tariff changes spanning 1840–2024. In doing so, this paper supersedes Den Besten and Känzig (2025) and Barnichon and Singh (2025), providing a unified and comprehensive narrative account of U.S. tariff policy over nearly two centuries and placing central emphasis on addressing the endogeneity of tariff changes when estimating their macroeconomic effects.

Finally, we relate to a large theoretical literature that studies the mechanisms and desirability of tariff policy. In international trade, a long tradition examines the long-run implications of tariffs and the structure of optimal tariff policies (e.g., Eaton and Kortum, 2002; Ossa, 2014; and more recently Caliendo, Kortum, and Parro, 2025; Itskhoki and Mukhin, 2025). In macroeconomics, several papers analyze the short-run dynamic effects of tariff shocks within open-economy general-equilibrium frameworks (e.g., Auclert, Rognlie, and Straub, 2025; Rodríguez-Clare, Ulate, and Vasquez, 2025). Another line of work studies the interaction between tariff shocks and monetary policy, highlighting how tariffs can generate inflationary pressures and alter optimal stabilization responses (Werning, Lorenzoni, and Guerrieri, 2025; Bergin and Corsetti, 2023; Monacelli, 2025; Bianchi and Coulibaly, 2025; Auray, Devereux, and Eyquem, 2025; Ambrosino, Chan, and Tenreyro, 2026). We contribute to this literature by providing empirical guidance on the aggregate effects of tariff shocks.

Outline. The paper proceeds as follows. In Section 2, we detail our newly assembled dataset spanning macroeconomic and financial information from 1840 at the annual and quarterly frequency. Section 3 provides a detailed historical account of U.S. tariff policy and describes our narrative approach to isolate plausibly exogenous tariff changes. In Section 4, we introduce our econometric approach. Section 5 reports the main results, discusses mechanisms and performs a series of sensitivity checks. Section 6 concludes.

2. Data and Measurement

Our empirical analysis relies on newly assembled U.S. macroeconomic data covering the period from 1840 until 2024 at both the annual and quarterly frequency. This section summarizes the main data sources and describes key measurement challenges, with particular attention to the construction of tariff variables.

Measuring the tariff rate. A central challenge for our analysis is how to quantify the stance of U.S. tariff policy over a long historical period. The literature typically relies on two revenue-based indicators: the average effective tariff rate (ATR) and the average tariff rate on dutiable imports, also known as the dutiable tariff rate (DTR):

$$\text{ATR}_t = \frac{\text{Duties Collected}_t}{\text{Total Imports}_t}, \quad \text{and} \quad \text{DTR}_t = \frac{\text{Duties Collected}_t}{\text{Dutiable Imports}_t}.$$

Although widely used in the trade and economic history literatures, these revenue-based measures have well-known limitations that complicate their use as indicators of the size of tariff changes (Irwin, 1998a). Three issues are particularly important in our long-run setting.

First, high-tariff goods are systematically underweighted. When duties are prohibitive, imports of the affected goods collapse, causing those items to drop out of the import base used to compute the average rate. As a result, average tariff measures tend to understate the true level of protection in periods of high or prohibitive tariffs.

Second, import-composition effects mechanically bias revenue-based measures downward. Tariff changes often induce substitution toward duty-free or low-tariff goods. Even if statutory schedules remain unchanged, such shifts lower the ratio of duties collected to imports. In principle, fixed-weight measures avoid this problem and typically show larger movements. In practice, however, fixed-weight series are not always readily available, especially for long historical periods.[1]

Third, the presence of specific duties introduces price-level sensitivity. When tariffs are levied as fixed charges per unit (e.g., $5 per ton of steel), their ad valorem equivalent declines when import prices rise and increases when prices fall. This creates an inverse, non-policy-related relationship between measured average tariff rates and import prices.

[1] Lerdau (1957) constructs a fixed-weight tariff index by converting duties to ad valorem rates and weighting them using 1949 wholesale price index expenditure shares. He finds that this measure is highly correlated with the tariff rate on dutiable imports, suggesting that while the latter does not fully eliminate composition effects, it provides a useful proxy for changes in overall tariff protection when fixed-weight data are unavailable.

In periods when specific duties were prevalent, fluctuations in import prices can therefore generate substantial variation in the average tariff rate that does not reflect underlying changes in trade policy.

In our analysis, we focus on the average tariff rate on dutiable imports, which is less affected by shifts in import composition and by the presence of duty-free goods. As pointed out by Irwin (1998a), duty-free imports typically lack close domestic substitutes; excluding them therefore yields a tariff measure that more accurately reflects the degree of protection granted to import-competing industries. Nonetheless, all revenue-based tariff measures contain potentially meaningful non-policy variation and therefore provide only a noisy proxy for the actual tariff policy stance. Beyond the inherent endogeneity of tariff rate changes, this noise further motivates the construction of an instrument for tariff shocks, allowing us to address both endogeneity and measurement-error concerns.

Annual data. To exploit rich historical variation in tariff policy, we construct a new annual dataset for the United States covering the period 1840–2024. We focus on annual variation, as these data are more consistently available and reliable over the long historical sample than higher-frequency measures.

Using historical publications, archival material, and modern statistical sources, we collect a wide range of macroeconomic and financial indicators, including measures of tariff protection, real economic activity, prices and wage dynamics, external balances, and financial variables such as interest and exchange rates. Details are provided in Appendix B, but we summarize the main features of the dataset here.

For 1840–1945, we obtain data on duties, imports, and exports from *Historical Statistics of the United States, 1789–1945* (U.S. Census Bureau, 1949). Until 1915, these series are reported on a fiscal-year basis; thereafter, they are reported on a calendar-year basis. To ensure consistency across variables, we convert fiscal-year observations to calendar years using our constructed quarterly series, detailed below. For later years, we use the *Historical Statistics of the United States: Earliest Times to the Present, Millennial Edition* (Carter et al., 2006), and from 1988 onward, data are from the United States International Trade Commission.

Real GDP, the GDP deflator, and the interest rate spread are taken from the Global Macro Database (Müller et al., 2025).[2] Labor compensation data are from Officer and Williamson (2025), which provides a long-run consistent series of average hourly compensation (earnings plus benefits) for production workers in manufacturing.

[2]In Appendix E.6 we assess the robustness of our results using alternative sources for real GDP and the GDP deflator, as well as industrial production in place of real GDP, and obtain similar estimates.

The manufacturing production index is constructed by combining several historical data sources. We start with the series from Frickey (1947), available from 1860, and splice it to the series from Fabricant (1940), which starts in 1899. From 1919 onward, we use the Federal Reserve's Industrial Production: Manufacturing index.

We construct the terms of trade index as the ratio of the export price index to the import price index, using historical merchandise price indices from North (1961), Simon (1960), and Lipsey (1963), combined with the BEA's export and import deflators from 1929 onward.

Finally, we construct a historical U.S. real effective exchange rate (REER) index using import-share trade weights for major U.S. trading partners, following the methodology outlined by the Bank for International Settlements (BIS). We compute trade weights in non-overlapping three-year blocks, using import data from the *Historical Statistics of the United States: Earliest Times to the Present, Millennial Edition* (Carter et al., 2006), chain-link the resulting indices to obtain a continuous series, and combine it with the BIS REER from 1964 onward. Reassuringly, our historical index closely matches the BIS narrow REER over the period they overlap.

Quarterly data. To improve the timing of tariff shocks and mitigate concerns that annual tariff changes may be confounded with other contemporaneous developments, we also construct a new quarterly dataset for the United States extending back to 1855. Quarterly tariff data for the 19th and early 20th centuries are not readily available in digitized form, so building a long-run quarterly series requires combining manual digitization of historical sources (including microfiche files and scans of official publications and reports).

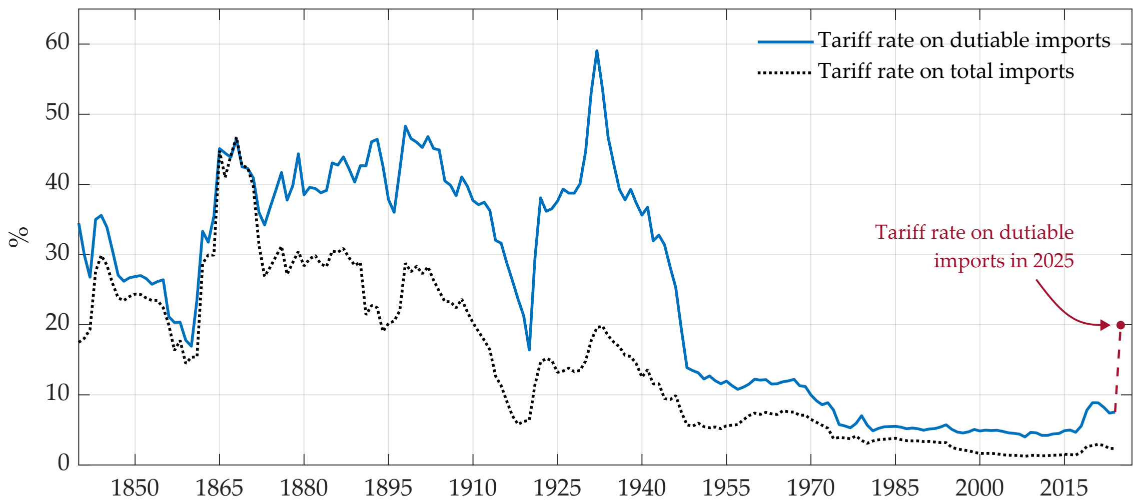
Because quarterly data are unavailable for many variables in the early sample, we follow Balke and Gordon (1986) and apply the Chow–Lin temporal disaggregation method to distribute annual observations across quarters using relevant high-frequency indicators, while preserving annual totals. As in Antolin-Diaz and Surico (2025), indicators are chosen on a variable-by-variable basis. The disaggregated series are spliced to observed quarterly data when available. This yields quarterly series for duties, imports, exports, output, prices, and other variables back to the mid-nineteenth century. Working at the quarterly frequency sharpens the timing of tariff shocks relative to macroeconomic outcomes, but entails a trade-off, as the underlying variables are generally measured more reliably at the annual frequency.

3. Identification

3.1. A History of U.S. Tariff Policy Since 1840

We begin with a brief historical overview of U.S. tariff policy since the mid-nineteenth century. Figure 1 plots the average U.S. tariff rate, measured as customs duties calculated over imports, from 1840 to 2024 and provides a visual summary of this history.[3]

Figure 1: Average U.S. Tariff Rates, 1840–2024



Notes: Average tariff rates in the U.S. from 1840 to 2024 (in percent). The blue line shows the rate expressed relative to dutiable imports; the black dotted line shows the rate relative to total imports.

Economic historians typically divide U.S. trade policy into three broad eras: a revenue period (roughly 1790–1860), when tariffs were primarily a source of government finance; a period of restriction (1861–1933), characterized by high and often volatile protection; and a reciprocity period (from 1934 onward), dominated by negotiated reductions in trade barriers (Irwin, 2020). Our analysis starts in the late revenue period and covers the full protection and reciprocity eras.

From the advent of the Second Party System through the early twentieth century, tariffs were at the core of partisan conflict. Democrats consistently advocated low tariffs "for revenue only", while Whigs and later Republicans defended high protective tariffs. This divide had a clear geographic foundation: Democrats drew most of their support from the agrarian, export-oriented South, whereas Republicans represented the industrial North, where manufacturers faced foreign competition (Irwin, 2017a).

[3] In Appendix B.3, we provide additional descriptive statistics showing that the United States has long been meaningfully exposed to international trade and that the composition of trade has evolved gradually over time, with a growing role for manufactured and consumer goods.

Mid-nineteenth century: protection vs. revenue. After a decade of gradually declining duties under the Compromise of 1833, the economy worsened in the late 1830s and the government faced growing budget deficits. This gave the Whig Party an opportunity to revive protective tariffs through the Tariff of 1842. Democrats criticized the tariff as unfair, arguing that it benefited capital more than labor, and instead supported a "revenue tariff with incidental protection". When Democrats regained power in 1844, they enacted the Walker Tariff of 1846, significantly lowering duties again. Continued economic growth and large federal surpluses led to a further reduction through the Tariff of 1857, which brought duties to their lowest levels in the nineteenth century (Irwin, 2017a).

The subsequent Panic of 1857 quickly reversed the political environment, as the recession turned fiscal surpluses into large deficits and created pressure to raise tariff revenue. Republicans were able to pass the Morrill Tariff in early 1861 after several Southern states seceded from the Union, resulting in the loss of a dozen Democratic senators who had previously blocked the higher tariffs. Within days, the Civil War began, and Congress enacted a sequence of further wartime increases that pushed average duties on dutiable imports well above prewar levels, inaugurating what Irwin terms the "protectionist era" of U.S. trade policy.

After the war, high tariffs persisted, both as a revenue source and as an instrument of industrial protection. From roughly 1870 to World War I, tariff law oscillated in response to partisan control. Republican administrations enacted or defended high protective schedules (e.g. the McKinley Tariff of 1890 and the Dingley Tariff of 1897), while Democratic victories were typically followed by reductions (e.g. the Wilson-Gorman Tariff of 1894). This sequence of tariff changes illustrates how similar fiscal and economic conditions led to opposite tariff responses depending on political majorities (see e.g. Irwin, 1998b). Even though Republicans occasionally also enacted modest tariff reductions, these were driven by strategic political considerations, primarily to preempt deeper Democratic cuts or to respond to public dissatisfaction with high duties (e.g., the Tariff Acts of 1870 and 1872), rather than by underlying economic pressures (Irwin, 2017a).

Early twentieth century: from progressive reform to Smoot-Hawley. At the turn of the century, critics argued that high tariffs protected trusts and raised consumer prices, fueling a split between Old Guard protectionists and progressive Republicans. In response, the Republican Party platform sought to put tariff policy on a "scientific" basis by advocating that tariffs should equalize the difference in production costs between the United States and foreign competitors (Dobson, 1976). The Payne-Aldrich Tariff (1909) modestly reduced average duties, largely as a compromise within the Republican coalition.

The election of Woodrow Wilson in 1912, aided by this intra-party split, led to the Underwood–Simmons Tariff (1913) – a major reduction in rates, framed explicitly as progressive reform and accompanied by the new federal income tax under the 16th Amendment. Contemporary accounts present Underwood-Simmons as the high-water mark of a broader effort to dismantle the Civil War-era protectionist regime, again driven primarily by ideology and distributional politics rather than short-run macroeconomic considerations (Irwin, 2017a; Brownlee, 2017).

The interwar period then brought a dramatic swing in the opposite direction. First, in response to the postwar recession, Congress enacted the 1921 Emergency Tariff in an effort to shield farmers from falling crop prices. The temporary measure was replaced by the Fordney–McCumber Act in 1922, which raised tariffs on industrial goods above their 1909 levels and further increased agricultural duties. The Smoot–Hawley Tariff Act of 1930 pushed rates even higher. While it originated as an effort by progressive Republicans to satisfy Midwestern farm interests, logrolling and sectoral lobbying transformed the bill into a tariff increase affecting many sectors, raising average tariffs on dutiable imports to their highest peacetime level (Irwin, 2017a). Although the Great Depression shaped the political context and amplified pressures from import-competing interests, the Smoot–Hawley bill was introduced before the 1929 stock market crash and originated in a distributional rationale, namely extending to agriculture the kind of tariff protection long associated with manufacturing. In Figure 1, this appears as the last major spike in average tariffs before the post-1934 downward trend.

The reciprocity era: bilateral agreements and multilateral liberalization. The Reciprocal Trade Agreements Act (RTAA) marked an important shift in the process of trade policymaking: it authorized the President to negotiate bilateral tariff reductions, and embedded U.S. trade policy in a framework of reciprocal concessions rather than unilateral protection (Irwin, 2017a). Congress specified the limits of this authority, including maximum permissible tariff reductions, and required periodic renewal. Within these constraints the President could implement negotiated tariff changes without additional legislative approval (CBO, 1979). Over time, this institutional innovation reshaped the political economy of trade by strengthening export-oriented interests and diluting the dominance of narrowly protectionist coalitions.

After World War II, U.S. tariffs declined steadily through successive rounds of the General Agreement on Tariffs and Trade (GATT) – Geneva (1947), Annecy (1949), Torquay (1950–51), Geneva II (1955–56), Dillon (1961–62), Kennedy (1964–67), and Tokyo (1973–79) – culminating in the Uruguay Round (1986–94), which created the World Trade Organization (WTO). These multilateral negotiations produced substantial cumulative reductions in industrial tariffs and extended rules to new areas such as agriculture, services, and intellectual property (Irwin, 2017a).

Politically, support for trade liberalization became increasingly bipartisan during the Cold War, linked to broader foreign-policy objectives – containing communism and fostering Western integration – and to an emerging consensus that export growth was central to U.S. prosperity. Tariffs in this era were rarely used as countercyclical tools. Two notable exceptions are the Nixon surcharge of 1971, which was motivated by balance-of-payments concerns, and the temporary oil import fees under Ford in 1975, which were a direct response to the 1973–74 oil shock and the subsequent recession and inflation.

The late twentieth century and the new protectionism. By the late twentieth century, statutory tariffs on industrial goods in the United States and other advanced economies were both low in level and relatively stable over time. Trade conflicts increasingly played out through non-tariff barriers, antidumping duties, and regional agreements such as the North American Free Trade Agreement (NAFTA), rather than large across-the-board changes in tariff schedules.

The Trump administration's tariffs in 2018–2019 and the subsequent 2025 tariff package represent a notable departure from the long postwar trend toward trade liberalization. These measures raised average U.S. tariffs from their post-Uruguay lows and, consistent with the administration's "America First" trade agenda, were justified on grounds including national security concerns, alleged "unfair" trade practices, and large bilateral trade imbalances. In Figure 1, they show up as modest but clearly visible upticks at the very end of the sample. Recent survey evidence suggests that, as in the late nineteenth century, attitudes toward tariffs have once again become more polarized along partisan lines, with Republicans disproportionately supporting new protection and Democrats largely opposing it (Dolan et al., 2025).

3.2. A New Series of Exogenous Tariff Shocks

From this narrative account of the history of tariffs in the United States, two features stand out. First, when considering a longer sample period, tariff levels display substantial variation. Before the 1940s, average tariffs were high – often in the 30–60 percent range on dutiable imports – and exhibited large, discrete movements around major reforms. Second, tariff changes were frequently enacted for ideological and distributive reasons, driven by a strong partisan divide on the issue, rather than as a response to short-run economic conditions. Motivated by these features, we perform a detailed narrative account of key tariff policy changes in the U.S. from 1840 until 2024. Our aim is to identify the timing, content, and stated motivations of the major tariff changes in our sample. We present the main insights from this analysis here; details are provided in our narrative companion paper (Den Besten and Känzig, 2026).

Sources and scope. Our analysis focuses on broad-based tariff changes that apply widely across products or sectors and therefore have the potential to affect aggregate macroeconomic conditions. We exclude narrow or highly targeted interventions, such as the 1983 motorcycle tariffs or the 2002 steel safeguards under President Bush, as these measures were limited in scope and unlikely to generate meaningful economy-wide effects. We also abstract from bilateral or regional trade arrangements, including the reciprocal agreements negotiated under the RTAA and later preferential trade agreements, since these typically produced incremental, partner-specific adjustments rather than discrete, uniform changes in statutory tariff rates. Our goal is to identify the major episodes of tariff reform that plausibly constitute aggregate policy shocks.

Our primary source for identifying the major U.S. tariff changes and understanding the motivations behind them is Irwin's comprehensive political and economic history of American trade policy (Irwin, 2017a). To ensure completeness and accuracy, we systematically cross-check the episodes discussed by Irwin with a wide range of additional historical sources and legislative records. These include classic tariff histories (e.g., Taussig, 1910; Fisk, 1910; Dobson, 1976; Tarbell, 1911), contemporaneous government communications (such as presidential speeches and official statements), statutory and archival documents (such as the Congressional Record, the United States Statutes at Large, and Federal Register notices for modern tariffs), political accounts (e.g., Flaherty, 2001; Pastor, 1980), and empirical studies of specific reforms (e.g., Eichengreen, 1999; Irwin and Kroszner, 1996). For the postwar period, our sources include detailed government reports and descriptive analyses of GATT negotiations (e.g., Irwin, 1995; CBO, 1979; GAO, 1994), as well as recent analyses of the Trump-era tariffs (Amiti, Redding, and Weinstein, 2019; Bown, 2021; Irwin, 2017b). This multi-source approach allows us to reconstruct a consistent narrative of major tariff reforms and their underlying motivations across nearly two centuries.

Endogenous vs. exogenous tariff changes. Our classification of tariff changes follows the narrative-identification logic of Romer and Romer (2010), who distinguish tax reforms taken for reasons unrelated to current or prospective macroeconomic conditions from those enacted in response to economic developments. We apply the same principle to trade policy. We classify a tariff change as plausibly exogenous when the narrative record indicates that its primary motivation was independent of contemporaneous or anticipated real economic activity. Such motivations include ideological commitments to protectionism or free trade; distributional concerns over the incidence of tariff policy; and political considerations, encompassing legislative strategy as well as diplomatic and geopolitical objectives.

By contrast, we classify a tariff change as endogenous when the historical record indicates that it was undertaken in response to economic conditions or fiscal pressures. These motivations include the need to raise revenue, for example to finance wartime expenditures or to offset cyclical declines in customs receipts; attempts to stabilize domestic economic activity by influencing inflation, output, or employment; and actions taken to address external imbalances, such as balance-of-payments pressures. In these cases, tariff policy is explicitly used as a tool of macroeconomic or public-finance management, and the resulting policy changes are therefore not suitable for identifying the effects of tariff shocks.

Based on the sources detailed above, we were able to identify 35 major tariff reforms from 1840 until 2024. Out of these, we classify 21 tariff changes as exogenous and 14 as endogenous. Appendix A lists all the events and provides information on their timing, the main motivation, and the event classification. Details on the narrative classification are provided in Den Besten and Känzig (2026).

To illustrate our approach, it is useful to highlight a few concrete examples. A first, exogenous episode is the McKinley Tariff of 1890. The narrative record indicates that the reform was primarily driven by an ideological, protectionist commitment. Republicans interpreted their electoral victory in 1888 as a mandate for strengthening tariff protection. McKinley's justification of the bill on the House floor reflected the belief that protective tariffs would best promote long-run national economic development (McKinley, 1890). There is no indication that the legislation was motivated by weak economic activity or fiscal stress; rather, it reflects a classic protectionist objective.

A second plausibly exogenous case is the Underwood–Simmons Tariff of 1913. Presidential speeches and House floor debates indicate that the tariff reduction was primarily motivated by Democratic ideological opposition to protectionism – specifically, the belief that freer trade would promote efficiency and fair competition – and by distributional concerns about tariffs benefiting large firms at consumers' expense, rather than by contemporaneous cyclical conditions.

In contrast, some major tariff changes are clearly endogenous to macroeconomic or fiscal developments. A salient example is the Revenue Act of 1861, and the sequence of Civil War-era increases that followed. These measures were explicitly adopted to finance wartime expenditures: the narrative sources describe them as revenue-raising instruments in the face of extraordinary spending needs, not as autonomous policy shifts. A second endogenous episode is the Emergency Tariff Act of 1921. Passed during the sharp postwar recession, the act raised duties on a range of agricultural imports with the stated goal of alleviating distress among farmers facing collapsing prices. The measure was framed as a temporary, emergency response to acute economic conditions and was explicitly justified on stabilization grounds. In both episodes, tariff policy was used instrumentally to address contemporaneous macroeconomic or fiscal challenges.

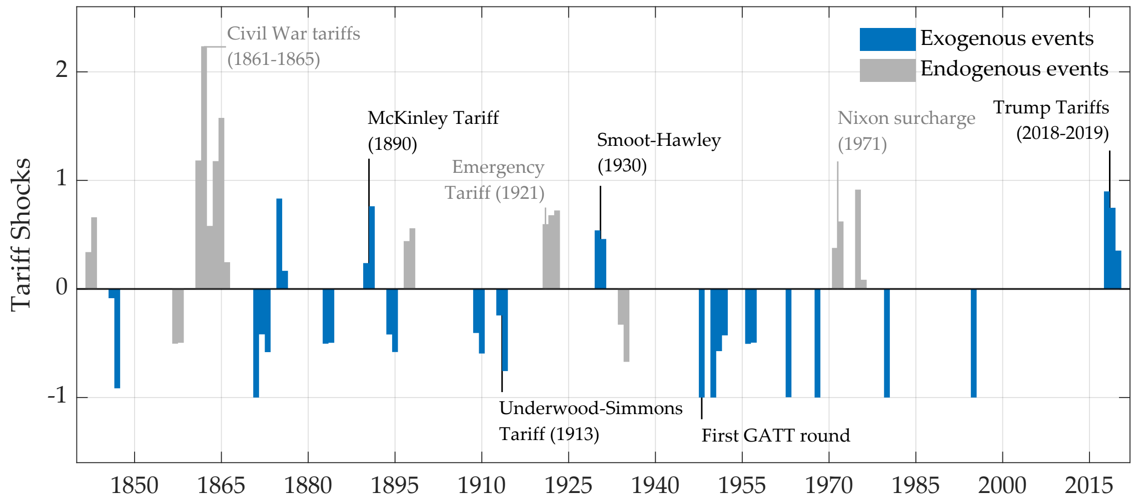
Measuring tariff shocks. In constructing our baseline shock series, we follow the spirit of Romer and Romer (1989) and focus on the timing and direction of identified tariff changes rather than on the precise magnitude of the rate adjustments.

For each episode classified as exogenous, we assign a tariff shock $\mathcal{T}_t \in \{-1, 1\}$ with 1 indicating tariff increases and -1 indicating tariff reductions. This sign-only approach offers two advantages. First, it provides a conservative measure of policy changes in settings where statutory or effective ad valorem rate changes may be difficult to quantify consistently across time – for example, when specific duties or mixed schedules complicate the mapping into comparable percentage changes. Second, and more importantly for identification, relying on the sign rather than the size of the reform helps isolate the component of the policy change that is plausibly exogenous. While the timing and direction of a reform may be plausibly exogenous – for example, because a shift in political power or a long-standing ideological objective created the opportunity for action – the magnitude of the associated tariff adjustment may still be shaped by contemporaneous economic or fiscal considerations. Focusing on the sign of the reform helps isolate the component of trade policy most likely to reflect an autonomous policy shift.

As the baseline, we thus construct an annual series that records only the timing and sign of each exogenous tariff episode. Under the assumption of exogeneity, such qualitative proxies identify the dynamic causal effects of interest (Boer and Lütkepohl, 2021; Plagborg-Møller, 2022). In robustness exercises, we consider alternative scalings based on changes in the average tariff rate (see Appendix E.2.3).

Because tariff reforms are implemented on specific calendar dates rather than at uniform intervals, we distribute each narrative shock across the two adjacent years in proportion to the number of days remaining in the year of implementation. Let d_t denote the number of days left in year t after the reform takes effect, and let D_t denote the total

Figure 2: Narrative Tariff Shock Series



Notes: Tariff shock series based on our narrative account of U.S. trade policy, distributed over years of policy influence. Blue bars: plausibly exogenous tariff changes. Gray bars: endogenous tariff changes.

number of days in that year. For a tariff reform occurring in year t, we assign

$$\tilde{\mathcal{T}}_t = \frac{d_t}{D_t}\mathcal{T}_t, \quad \text{and} \quad \tilde{\mathcal{T}}_{t+1} = \left(1 - \frac{d_t}{D_t}\right)\mathcal{T}_t,$$

where $\mathcal{T}_t \in \{-1, 1\}$ is the sign of the identified tariff change. This timing convention avoids placing the full weight of a policy change on the exact enactment date and mitigates the influence of reporting granularity, while also accommodating the implementation lags that are common in tariff reforms.

Figure 2 shows the resulting narrative tariff shock series. Over the full sample, we identify six exogenous tariff increases and fifteen exogenous reductions.[4] Tariff increases occur primarily in the earlier part of the sample, whereas the postwar period is characterized predominantly by exogenous tariff cuts.

Diagnostics. We perform a series of diagnostic checks on the narrative shock series (Ramey, 2016). We find little evidence of serial correlation and no evidence that macroeconomic or financial variables – such as output, exports, imports, the terms of trade, stock prices, or government finances – predict the shock series. Table 1 reports the p-values for the corresponding Granger causality tests. For all variables considered, the p-values

[4] We exclude the exogenous Fordney–McCumber Tariff Act in our annual specification because it partly overlaps with the endogenous Emergency Tariff due to the distributed value, leaving effectively five tariff increases. Our results are robust to a more balanced specification that includes the Fordney–McCumber Tariff increase as well as the 1964 Chicken Tax, which is less broad-based and therefore not part of our baseline set of events, see Appendix E.2.4.

are well above conventional significance levels, with a joint p-value of 0.91, indicating that the shocks are not systematically anticipated by observable economic conditions. By contrast, the same variables are strong predictors for endogenous tariff events, consistent with those actions being taken in response to underlying macroeconomic developments.

Table 1: Forecastability of Exogenous and Endogenous Tariff Shocks

	Narrative tariff shocks	
Variable/p-value	Exogenous	Endogenous
Tariffs and trade		
Instrument	0.13	0.00
Tariff rate	0.92	0.11
Terms of trade	0.93	0.36
Real exports	0.11	0.03
Real imports	0.16	0.26
Macroeconomic variables		
Real GDP	0.29	0.10
Manufacturing production	0.33	0.14
Real compensation	0.63	0.91
GDP deflator	0.75	0.01
Producer price index	0.41	0.47
Financial variables		
Interest rate spread	0.94	0.04
REER	0.54	0.88
Real energy commodity price index	0.53	0.26
SP500	0.59	0.01
Government finances		
Real government debt	0.86	0.46
Real government revenue	0.69	0.02
Real government expenditure	0.27	0.92
Joint	0.91	0.00

Notes: p-values of Granger causality tests of the narrative tariff shock series using a selection of macroeconomic and financial variables. Lag order: 2. Variables are transformed to be stationary before estimation.

To further assess whether economic conditions might indirectly influence tariff policy through political transitions, we investigate the predictability of unified governments in Appendix C. In our narrative analysis, our exogeneity classification often relies on the premise that tariff changes are ideologically driven and typically implemented when one party secures unified control. Reassuringly, we find no evidence that macroeconomic variables or recession episodes affect transitions to unified control, reducing concerns that economic conditions indirectly influenced the timing of tariff reforms.[5]

Finally, we show that the series is uncorrelated with other shock measures from the literature, including other fiscal and monetary policy shock measures, see Appendix C.

3.3. A First Look at the Data

To illustrate the identifying variation in our narrative tariff shocks, we begin with a simple event-study specification that relates short- and medium-run changes in macroeconomic outcomes to our narrative tariff shock. The goal is to offer a first visual summary of the relationship in the data. Specifically, we estimate:

$$y_{d+h} - y_{d-1} = \alpha_h + \beta_h \Delta \mathrm{DTR}_d + \rho_h \Delta y_{d-1} + \phi_h \Delta \mathrm{DTR}_{d-1} + \gamma_h' X_d + \nu_{d+h}, \tag{1}$$

where d indexes event years, y_{d+h} is the outcome of interest h years ahead, and $\Delta\mathrm{DTR}_d$ is the change in the dutiable tariff rate, instrumented with our narrative tariff shock $\tilde{\mathcal{T}}_d$. As outcome variables, we consider real GDP and the GDP deflator (in logs). We control for one lag of the outcome variable and the dutiable tariff rate, respectively. In addition, we include a set of deterministic controls – such as indicator variables for major historical episodes, including the World Wars – as detailed in Section 4.1. We report robust standard errors following Montiel Olea and Plagborg-Møller (2020).

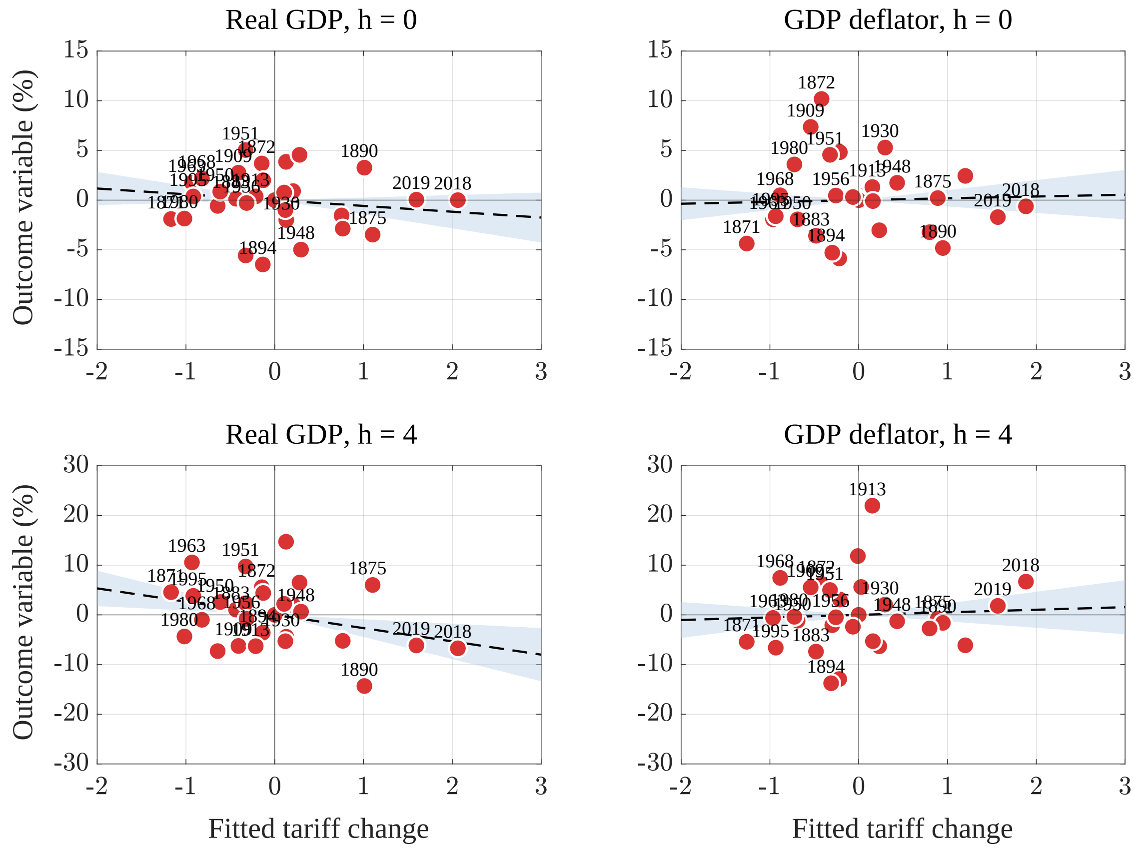
Figure 3 presents scatter plots of changes in log real GDP and the deflator, on impact and four years ahead, against the instrumented change in the dutiable tariff rate – both residualized with respect to the controls – along with the corresponding fitted line.

Two patterns emerge. First, exogenous tariff increases are associated with a decline in real activity. On impact, the estimated relationship between tariff changes and real GDP is negative but not statistically significant at the 90 percent level. At a four-year horizon, the relationship becomes considerably stronger: episodes of tariff increases are associated with sizable cumulative declines in real GDP, while tariff reductions coincide with sustained expansions. The negative relationship is robust and not driven by a small number of outliers. Second, the price response is imprecisely estimated and not statistically distinguishable from zero. The estimated relationship between tariff changes and the GDP deflator is slightly positive, but remains small and imprecise.

Overall, this event-study evidence provides a first visual summary of the raw rela-

[5]Unified control of government denotes periods in which the same political party controls the Presidency and holds majorities in both chambers of the U.S. Congress. Such unified government was often a necessary condition for the passage of tariff legislation; however, we show in Appendix E.2.1 that our results are not driven by such broader partisan shifts.

Figure 3: Narrative Tariff Shocks and the Economy



Notes: Scatter plot illustrating the relationship between the instrumented tariff rate and the cumulative change in outcome variables at horizons $h = 0, 4$, based on the event-study specification (1). As outcome variables, we consider the change in real GDP and the GDP deflator, in the current year and over a four year horizon. Both variables are residualized with respect to their lags and our deterministic controls. The sample includes only years in which the distributed narrative instrument records a tariff shock. Lag order: 1. Shaded areas represent 90 percent confidence bands based on heteroskedasticity-robust standard errors.

tionship between tariff changes and macroeconomic outcomes, based on a minimal set of controls. The results point to a robust negative relationship between tariffs and real activity that is not driven by outliers, while the price response does not exhibit a clear pattern. In the next section, we formalize the analysis in a semi-structural framework that exploits the full time-series variation to trace the dynamic causal effects of tariff shocks.

4. Empirical Approach

As illustrated above, our narrative tariff shock series has many desirable properties. Nonetheless, it is only an imperfect measure of the shock of interest because it likely does not capture all relevant tariff changes and could be measured with error (Stock and Watson, 2018). Therefore, we do not use it as a direct shock measure but as an *instru-*

ment. Provided that the narrative tariff shock measure is correlated with the structural tariff shock but uncorrelated with all other shocks, we can use it to estimate the dynamic causal effects of tariff shocks.

4.1. Framework

We are interested in modeling the U.S. economy. Let $\mathbf{y}_t$ denote a $n \times 1$ vector of annual time series. We assume that the dynamics of $\mathbf{y}_t$ can be characterized by the following structural vector moving-average representation:

$$\mathbf{y}_t = \mathbf{B}(L)\mathbf{S}\boldsymbol{\varepsilon}_t, \tag{2}$$

where $\boldsymbol{\varepsilon}_t$ is a vector of mutually uncorrelated structural shocks driving the economy, $\mathbf{B}(L) \equiv \mathbf{I} + \mathbf{B}_1 L + \mathbf{B}_2 L^2 + \ldots$ is a matrix lag polynomial, and $\mathbf{S}$ is the structural impact matrix. Assuming that the vector-moving average process (2) is invertible, it admits the following vector autoregression (VAR) representation:

$$\mathbf{A}(L)\mathbf{y}_t = \mathbf{S}\boldsymbol{\varepsilon}_t = \mathbf{u}_t, \tag{3}$$

where $\mathbf{u}_t$ is a $n \times 1$ vector of reduced-form innovations with variance-covariance matrix $\text{Var}(\mathbf{u}_t) = \boldsymbol{\Sigma}$ and $\mathbf{A}(L) \equiv \mathbf{I} - \mathbf{A}_1 L - \ldots$ is a matrix lag polynomial. Truncating the VAR to order p, we can estimate the model using standard techniques and recover an estimate of $\mathbf{A}(L)$.

We want to identify the causal impact of a single shock: the tariff shock. Without loss of generality, let us denote the tariff shock as the first shock in the VAR, $\varepsilon_{1,t}$. Our aim is to identify the structural impact vector $\mathbf{s}_1$, which corresponds to the first column of $\mathbf{S}$.

External instrument approach. Identification using external instruments works as follows. Suppose there is an external instrument available, z_t. In our application, the instrument is the narrative tariff shock series $z_t = \tilde{\mathcal{T}}_t$. For z_t to be a valid instrument:

$$\mathbb{E}[z_t \varepsilon_{1,t}] = \alpha \neq 0 \tag{4}$$

$$\mathbb{E}[z_t \boldsymbol{\varepsilon}_{2:n,t}] = \mathbf{0}, \tag{5}$$

where $\varepsilon_{1,t}$ is the tariff shock and $\boldsymbol{\varepsilon}_{2:n,t}$ is a $(n-1) \times 1$ vector consisting of the other structural shocks. Assumption (4) is the relevance requirement and assumption (5) is the exogeneity condition. Under these assumptions, the structural impact vector $\mathbf{s}_1$ is identified up to sign and scale:

$$\mathbf{s}_1 \equiv \boldsymbol{\theta}_0 \propto \frac{\mathbb{E}[z_t \mathbf{u}_t]}{\mathbb{E}[z_t \mathrm{u}_{1,t}]} \tag{6}$$

provided that $E[z_t \mathrm{u}_{1,t}] \neq 0$. Note that identifying the relative impact effects does not require invertibility; it is only needed for dynamic effects beyond impact.[6]

To facilitate interpretation, we scale the structural impact vector such that the shock corresponds to a 1 percentage point increase in the tariff rate. We implement the estimator with a 2SLS procedure and estimate the coefficients above by regressing $\hat{\mathbf{u}}_t$ on $\hat{\mathrm{u}}_{1,t}$ using z_t as the instrument. To conduct inference, we employ a residual-based moving block bootstrap (Jentsch and Lunsford, 2019).

Relaxing VAR assumptions. The VAR approach improves precision, allowing for sharper inference. However, it relies on two potentially restrictive assumptions. The first is invertibility, meaning that the model incorporates all relevant information needed to recover the structural shocks of interest. The second pertains to the dynamic VAR structure, with the key assumption being that a finite-order VAR adequately approximates the dynamics of the data generating process well. We perform two exercises to assess how restrictive these assumptions are.

First, we alternatively estimate the responses using the internal instrument approach (Plagborg-Møller and Wolf, 2019). This approach does not rely on invertibility but instead assumes that the instrument is orthogonal to leads and lags of the structural shocks:

$$\mathbb{E}[z_t \boldsymbol{\varepsilon}_{t+j}] = \mathbf{0}, \quad \text{for } j \neq 0. \tag{7}$$

Together with the exogeneity and relevance requirement (4)-(5), this identifies the dynamic causal effects of interest. The approach can be implemented by augmenting the VAR with the instrument ordered first, $\bar{\mathbf{y}}_t = (z_t,\ \mathbf{y}_t')'$, and computing the impulse responses to the first orthogonalized innovation, $\bar{\mathbf{s}}_1 = [\mathrm{chol}(\bar{\boldsymbol{\Sigma}})]_{\cdot,1} / [\mathrm{chol}(\bar{\boldsymbol{\Sigma}})]_{1,1}$.

Second, we estimate the impulse responses to the tariff shock using local projections. This approach relies on the same identifying assumptions as the internal instrument approach but directly estimates the responses without relying on the VAR structure, making it less prone to lag truncation bias. Specifically, we use a local projections-instrumental

[6] To be more precise, the VAR does not have to be fully invertible for identification with external instruments: it suffices if the shock of interest is invertible in combination with a limited lead-lag exogeneity condition (Miranda-Agrippino and Ricco, 2018).

variable approach (Jordà, Schularick, and Taylor, 2015; Ramey, 2016), directly estimating the dynamic causal effects of a tariff shock from the following set of regressions:

$$y_{i,t+h} - y_{i,t-1} = \alpha_h^i + \theta_h^i \Delta y_{1,t} + \boldsymbol{\beta}_h^{i\prime} \Delta \mathbf{x}_{t-1} + \nu_{i,t,h}, \tag{8}$$

using z_t as an instrument for $\Delta y_{1,t}$. Here, $y_{i,t+h}$ is the outcome variable of interest, $y_{1,t}$ is the tariff rate, $\mathbf{x}_{t-1}$ is a vector of controls, $\nu_{i,t,h}$ is a potentially serially correlated error term, and h is the impulse response horizon. We use the same controls as in the VAR. For inference, we follow the lag-augmentation approach proposed by Montiel Olea and Plagborg-Møller (2020).

Empirical specification. Our goal is to model the key channels through which tariff shocks affect the U.S. economy. Our baseline specification includes ten variables. To capture international and trade conditions, we include the average dutiable tariff rate, imports, exports, the terms of trade, and the real effective exchange rate. To measure real activity, we include real GDP and manufacturing production. To capture price and wage dynamics, we include the GDP deflator and the compensation of production workers. Finally, to account for financial conditions, we include a long-short interest rate spread. All variables enter in log levels, except for the tariff rate and the interest rate spread.

The main sample runs from 1866 to 2024. We begin after the Civil War to avoid measurement concerns affecting national accounts and trade statistics during the war years. We include two lags. Given the length of our sample, it is important to account for structural breaks and major confounding episodes. Because there is a clear shift in the level of tariff rates after World War II, we include a post-World War II dummy. We also include dummies for World War I, World War II, and the Great Depression. However, our results turn out to be robust to these choices. Further details on data sources and construction are provided in Appendix B.

Estimating effects on additional outcome variables. To analyze a broader set of outcome variables, we adopt a marginal VAR approach. Specifically, we augment the baseline VAR by adding one additional variable at a time and trace its impulse response, following Gertler and Karadi (2015).

5. Results

5.1. The Macroeconomic Effects of Tariffs

How do tariffs affect the macroeconomy? We now turn to this question by studying the dynamic effects of tariff shocks through the lens of our baseline VAR model, using the narrative tariff shock series as an instrument for the structural tariff shock.

The main identifying assumption behind the external instrument approach is that the instrument is correlated with the structural shock of interest but uncorrelated with all other structural shocks driving the economy. However, for standard inference, the instrument must also be sufficiently strong. To assess this, we conduct the weak instrument test by Montiel Olea and Pflueger (2013). The results indicate that the narrative tariff shock series is a strong instrument. The heteroskedasticity-robust F-statistic is 16.95 and thus exceeds conventional critical values. Overall, this evidence suggests that there is no weak instrument problem at hand and we proceed by conducting standard inference.

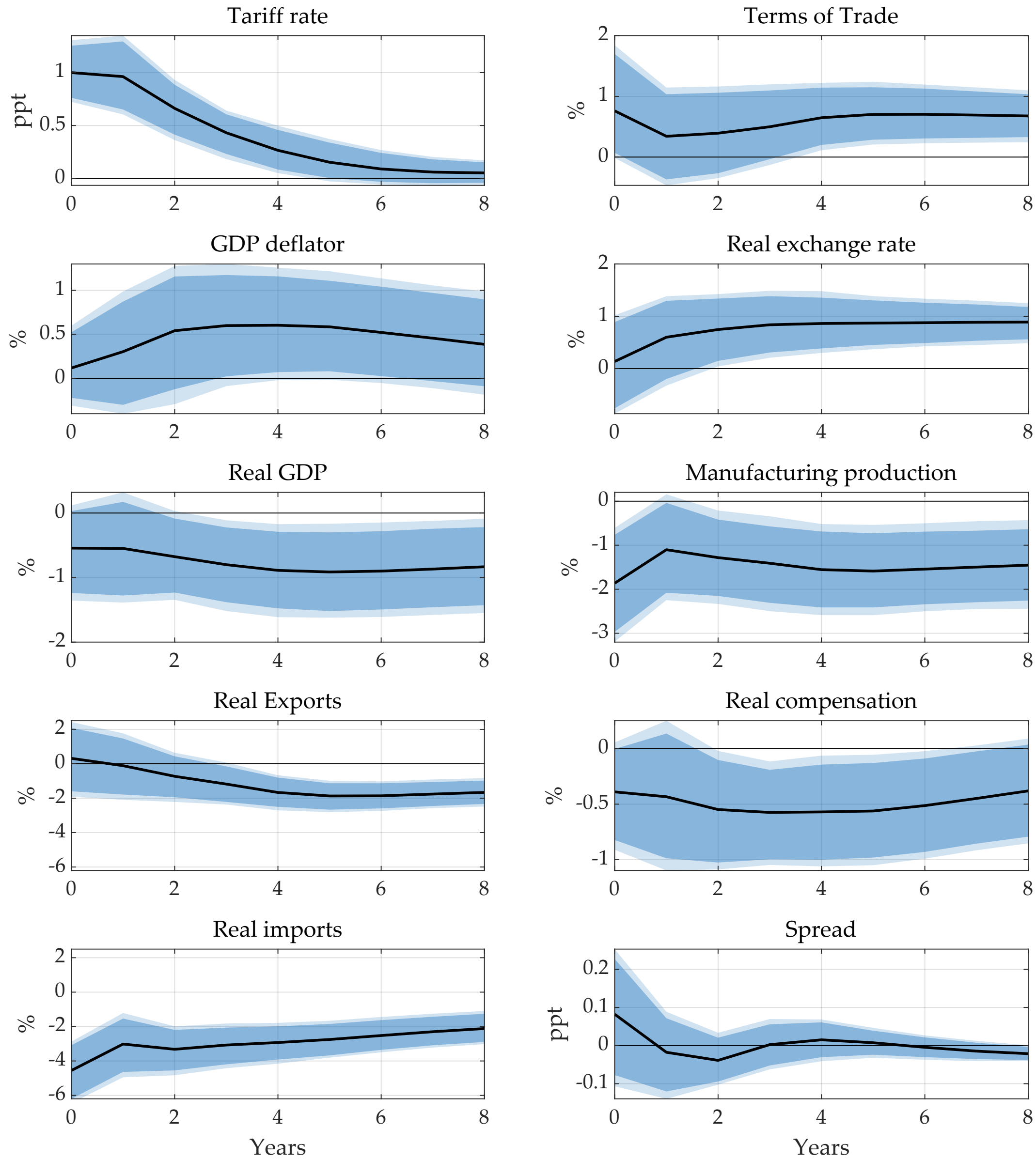
Figure 4 presents the impulse responses to the identified tariff shock. Throughout the paper, we normalize the responses such that the average tariff rate on dutiable imports increases by 1 percentage point on impact. This corresponds to a typical shock, amounting to roughly half a standard deviation. The solid black lines are the point estimates, while the shaded areas represent 90 and 95 percent confidence bands, respectively, computed from 2,000 bootstrap replications.

Following a tariff shock, the average tariff rate increases persistently and remains elevated for more than four years. This persistence is consistent with the historical record: many of the tariff changes identified in our narrative analysis were intended to be long-lasting and, in practice, proved difficult to reverse in the absence of the required political majorities. At the same time, the average tariff rate is an equilibrium object and may gradually adjust as firms and consumers respond to higher trade barriers by altering import composition or sourcing patterns. Despite these behavioral responses, tariff shocks lead to a persistent increase in tariff rates, with long-lasting effects on the trade environment.

The persistent increase in trade protection translates into immediate and substantial adjustments in trade flows. Imports contract sharply following a tariff shock, falling by roughly 4 percent on impact. The response is both large and rapid, reflecting the direct effect of higher trade barriers on import demand. As tariffs raise the relative price of foreign goods, firms and consumers adjust by reducing import volumes and shifting expenditure toward domestic or alternative sources.

Exports respond more gradually to a tariff shock. On impact, export volumes change

Figure 4: The Macroeconomic Effects of a Tariff Shock



First stage regression: robust F-statistic: 16.95, R^2: 6.21%

Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated based on the VAR model (3) using the narrative tariff shock series as an instrument. Lag order: 2. Solid line: point estimate. Dark and light shaded areas: 90 and 95 percent confidence bands based on moving-block bootstrap.

little and the effect is not statistically significant. Over time, however, exports fall, reaching a maximum decline of about 2 percent. A key mechanism behind this pattern is the sluggish appreciation of the U.S. dollar following the tariff shock, which erodes the international competitiveness of U.S. goods. In addition, the delayed response is consistent with the gradual adjustment of foreign demand and sourcing patterns.

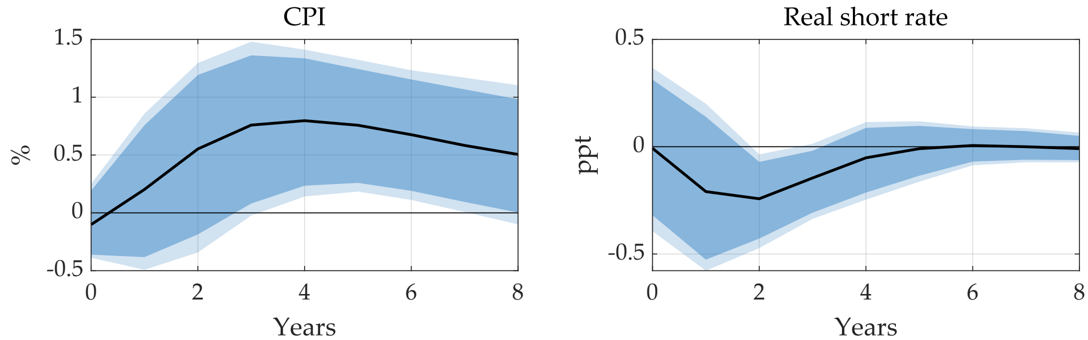
Together, the import and export responses imply a sizable contraction in overall trade activity. While tariffs are effective at compressing imports, they also weaken export performance through exchange rate adjustment and broader macroeconomic channels.

The contraction in trade is accompanied by a sizable and persistent decline in real economic activity. A tariff increase that raises the average tariff rate by 1 percentage point leads to a peak decline in real GDP of roughly 0.9 percent, with output remaining below trend eight years out. Interestingly, manufacturing production also decreases, by over 1.5 percent at peak. These estimates suggest that tariffs do not succeed at protecting domestic industry in the aggregate, in line with the findings in Steinberg (2025). Instead, higher trade barriers coincide with a broad-based weakening of industrial activity, underscoring the importance of general-equilibrium effects that offset any partial-equilibrium gains from reduced import competition. We also document a sustained decline in the compensation of production workers. These findings complement historical evidence showing that even when protection succeeded in fostering specific industries, it often came at substantial economy-wide costs (Irwin, 2000).

The terms of trade improve following a tariff shock, consistent with standard trade theory. Part of this improvement reflects a valuation effect: the appreciation of the domestic currency mechanically lowers import prices in domestic currency relative to export prices. The importance of this channel depends on currency invoicing: to the extent that imports are priced in domestic currency (e.g., dollar pricing), this valuation effect is muted or absent. More importantly, higher tariffs compress import demand, inducing foreign exporters to absorb part of the adjustment through lower prices, so that import prices fall relative to export prices (Johnson, 1953; Irwin, 1998a). This pattern is consistent with the classic optimal-tariff channel. However, despite this relative price gain, tariff shocks are contractionary overall: the favorable terms-of-trade movements do not offset the broader negative general-equilibrium effects on output and trade.

The response of aggregate prices is more nuanced. The GDP deflator increases modestly – by about 0.5 percent – but the estimate is imprecise, and the effect gradually dissipates as economic activity contracts. This pattern reflects the interaction of opposing forces. On the one hand, higher tariffs raise the cost of imported goods and intermediate inputs, exerting upward pressure on prices. On the other hand, the persistent decline in output, trade, and manufacturing activity generates demand-side disinflationary pressures that dampen the initial cost-push effect. As a result, tariffs do not translate into

Figure 5: The Effect on Consumer Prices and Interest Rates



Notes: Impulse responses of consumer prices and short-term interest rates to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated using the marginal VAR approach. Solid line: point estimate. Dark and light shaded areas: 90 and 95 percent confidence bands based on moving-block bootstrap.

sustained inflation at the aggregate level.[7]

Figure 5 shows additional responses of consumer prices and short-term interest rates to the tariff shock. The response of consumer prices is somewhat more pronounced than that of the GDP deflator. This pattern is consistent with the fact that the CPI includes imported goods, whereas the GDP deflator is based on domestically produced output. The real short-term interest rate changes little on impact and then falls modestly before gradually returning toward its pre-shock level. Longer-term rates display a similar pattern, as reflected in the statistically insignificant response of the long-short interest rate spread in Figure 4. Given the long sample and the substantial changes in monetary regimes over time – notably the gold standard and its eventual abandonment – this interest-rate response should be interpreted as an average historical pattern. We return to the role of monetary policy and its interaction with tariff shocks in more detail in Section 5.4.

The qualitative responses we document are broadly consistent with the predictions of modern trade and open-economy macro models, in which tariff increases compress imports, improve the terms of trade, and propagate through both supply- and demand-side channels (e.g. Caliendo, Kortum, and Parro, 2025; Auclert, Rognlie, and Straub, 2025). At the same time, the estimated magnitudes of the output, manufacturing, and trade responses tend to lie toward the upper end of the range implied by quantitative models, underscoring the importance of general-equilibrium spillovers through prices and wages. However, they are broadly consistent with the semi-elasticities estimated from narrative

[7]The importance of these demand-side channels is consistent with Barnichon and Singh (2025), although in our setting they attenuate rather than overturn the short-run inflationary response, with the difference likely reflecting unaccounted residual endogeneity in their tariff shock series (see Appendix D.3 for details).

tax shocks (Romer and Romer, 2010; Mertens and Ravn, 2013). For instance, Mertens and Ravn (2013) find that a 1 percentage point increase in the average personal tax rate lowers GDP by 1.8 percent at peak while a 1 percentage point increase in the average corporate tax rate reduces GDP by about 0.6 percent.[8]

Another notable feature of the estimated responses is the pronounced persistence of the effects of tariffs on GDP and manufacturing activity. This accords well with the idea that tariffs can depress innovation and slow productivity growth, as emphasized in recent work (Fornaro and Wolf, 2025), generating long-lived declines in economic activity.

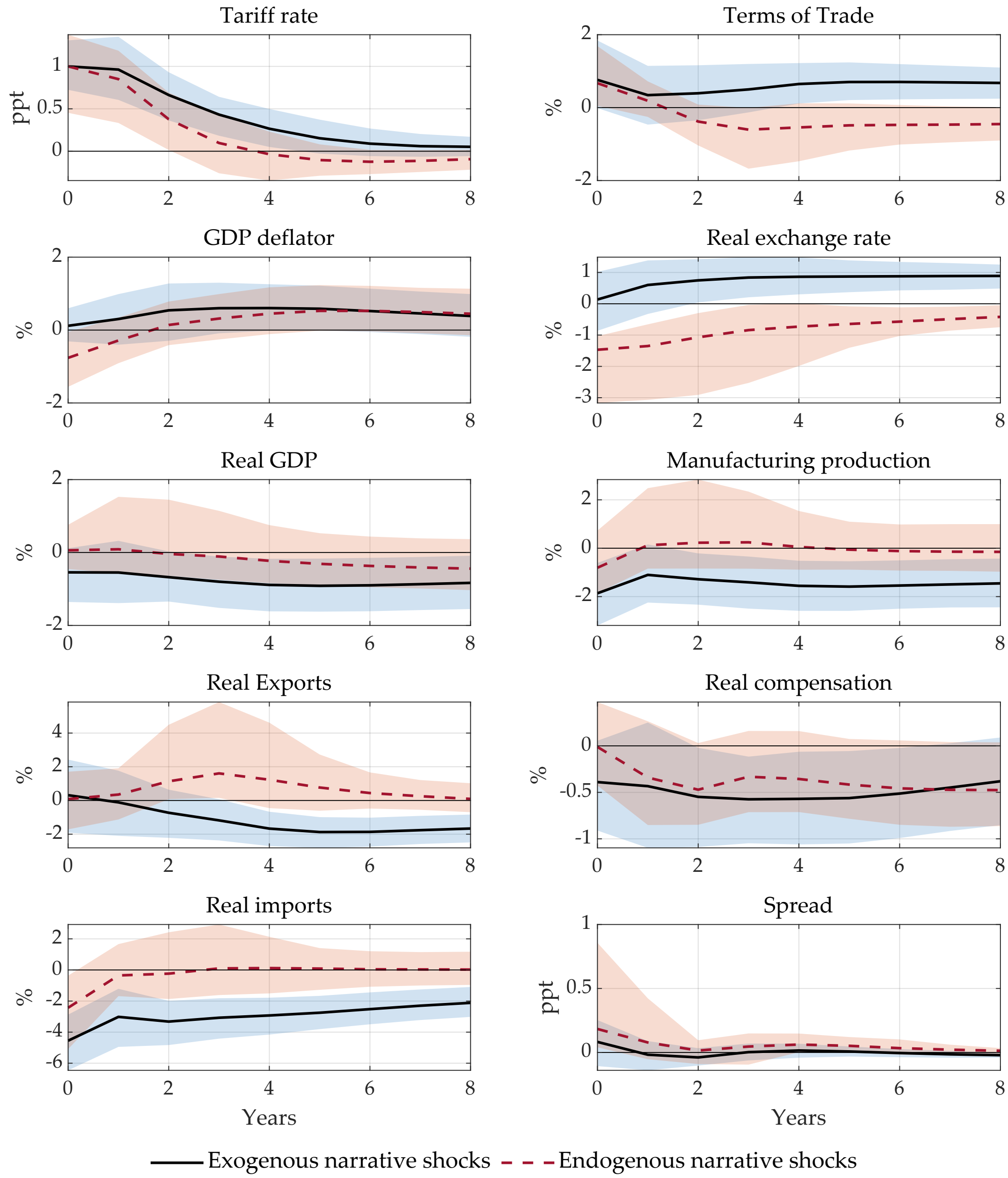
Taken together, these results paint a consistent picture of tariff shocks as persistent and contractionary disturbances. While tariffs sharply reduce imports by design, they also have broad-based macroeconomic consequences including a sustained decline in output, manufacturing activity, and exports, alongside more muted and transitory price increases. Rather than reallocating activity toward domestic production, the persistent rise in tariff protection translates into a lasting reduction in trade openness and a weakening of industrial performance in the aggregate. In this sense, tariffs fail to deliver on their stated objective of onshoring production once general-equilibrium effects are taken into account – highlighting the costs of protectionist policies relative to their intended benefits. Viewed through the same lens, our estimates also imply that reductions in tariff protection operate in the opposite direction, supporting trade, economic activity, and industrial output.

5.2. The Importance of Accounting for Endogeneity in Tariff Changes

A key motivation for our narrative identification strategy is the concern that tariff changes can be endogenous to economic conditions. Historically, many tariff reforms were indeed implemented in response to spending needs or the business cycle. In this section, we illustrate the importance of removing endogenous tariff changes by comparing our baseline results to an alternative instrument constructed from *endogenous* tariff events, that is from events that we classified as endogenous in our narrative identification step. We will see that failing to account for the endogeneity of tariff changes can lead to misleading inferences about their macroeconomic effects.

[8] Another way to compare the effects of tariffs to other taxes would be based on tax multipliers. However, this is challenging because tariff revenue is endogenous and depends on the response of imports. Depending on the level of tariffs and the elasticity of import demand, higher tariffs can even reduce, rather than increase, customs revenue (Irwin, 1998b). In our setting, we find a positive, but only modest and imprecise, response of revenues, rendering the tariff multiplier ill defined. A similar issue arises for corporate tax shocks in Mertens and Ravn (2013).

Figure 6: Exogenous vs. Endogenous Tariff Shocks



Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, using exogenous and endogenous narrative tariff shocks as the instrument, respectively. Lag order: 2. Solid line: point estimate. Dark and light shaded areas: 90 and 95 percent confidence bands based on moving-block bootstrap.

Figure 6 contrasts impulse responses obtained using our baseline instrument – based on plausibly exogenous tariff events – with those based on endogenous tariff changes. The differences are stark. When tariff changes are endogenous, imports and exports exhibit little systematic response, output shows no significant decline, the exchange rate depreciates rather than appreciates, and prices display a more pronounced disinflationary pattern.

These differences are informative. Endogenous tariff changes are enacted in response to prevailing macroeconomic and fiscal conditions and therefore tend to reflect those underlying developments rather than the causal effects of tariff policy itself. Consistent with this, Appendix Figure D.1 shows that endogenous tariff events are followed by a significant increase in government spending, in line with historical accounts emphasizing revenue-raising or fiscally motivated tariff changes. By contrast, there is no comparable response of government spending following plausibly exogenous tariff shocks. Taken together, these findings illustrate how failing to account for endogeneity in tariff changes can bias estimated responses.

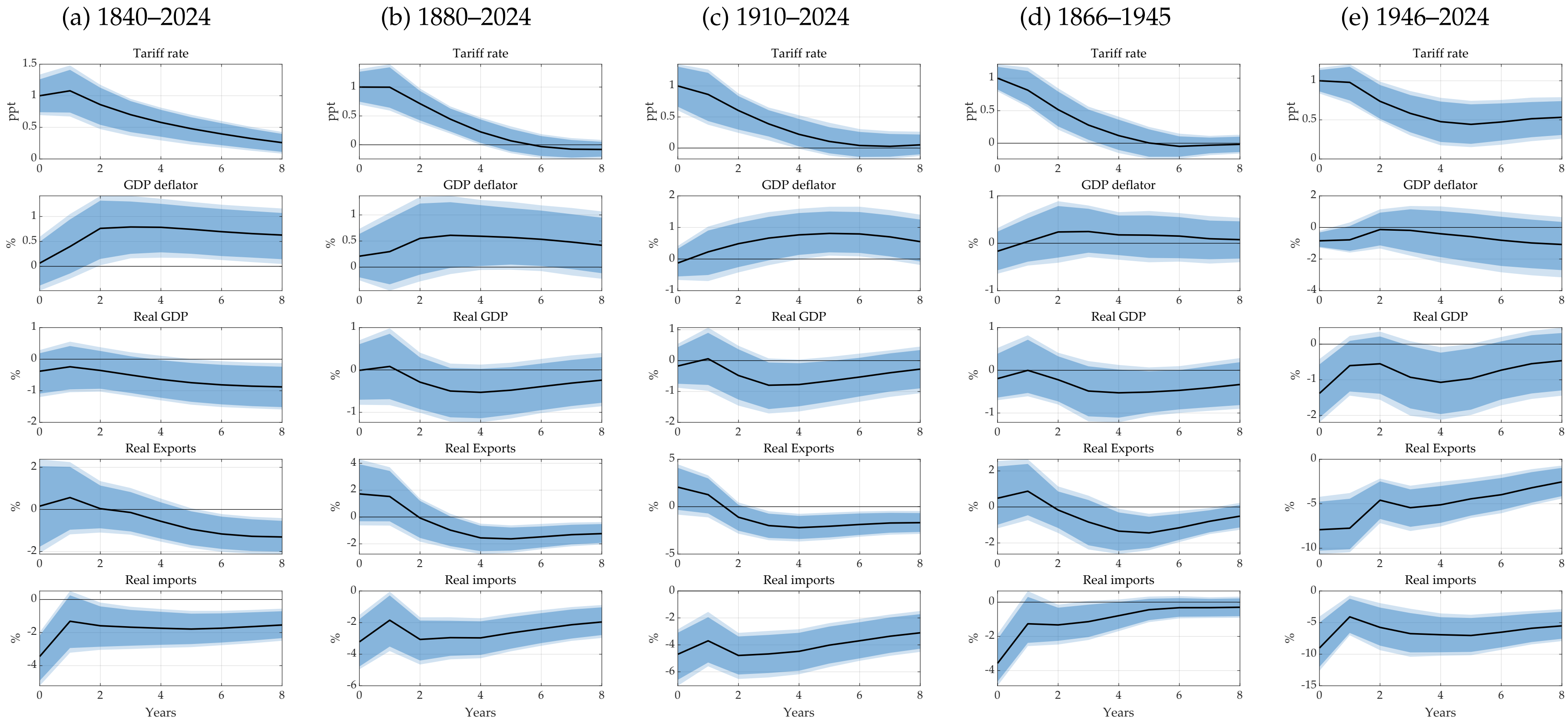
A careful narrative classification is therefore essential to isolate exogenous tariff shocks and obtain credible estimates of their causal impact. Consistent with this interpretation, Appendix D.2 further shows that a standard recursive identification scheme – which mixes endogenous and exogenous tariff changes – also yields responses that differ markedly from our baseline.

5.3. The Propagation of Tariff Shocks Over Time

Given the long historical sample period, structural relationships may have evolved over time. To assess the stability of our results over time, we re-estimate the baseline VAR on a series of subsamples spanning different institutional and economic regimes. Figure 7 reports impulse responses to a tariff shock normalized to increase the dutiable tariff rate by one percentage point, estimated over five samples: our full historical sample (1840–2024), a sample that excludes the reconstruction era (1880–2024), a sample starting in 1910 reflecting the modern fiscal era, a pre-World War II sample (1866–1945), and a postwar sample (1946–2024).

Across the first four subsamples, the responses are remarkably stable, both qualitatively and quantitatively. In each case, tariff shocks generate a persistent increase in tariff rates, a sharp contraction in imports, a delayed decline in exports, a sustained fall in real GDP, and a sluggish increase in prices. These patterns closely mirror our baseline estimates over the period 1866-2024. This stability is notable given the substantial changes in the structure of the U.S. economy, the evolution of fiscal institutions, and shifts in the international trading system over this long horizon. Taken together, the results suggest that the stagflationary effects of tariff shocks are not driven by a particular historical episode.

Figure 7: Macroeconomic Effects of Tariff Shocks Across Subsamples



Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated based on the external-instrument VAR model (3), over five different subsamples: 1840–2024, 1880–2024, 1910–2024, 1866–1945, and 1946–2024. Lag order: 2. Solid line: point estimate; dark and light shaded areas: 90 and 95 percent confidence bands based on moving-block bootstrap.

The main departure from this pattern emerges when comparing the pre-World War II and post-World War II samples. In the postwar period, tariff shocks generate a significantly different price and export response. Most notably, prices now fall persistently following a tariff shock, and exports decline sharply on impact rather than with a delay. Reflecting this more immediate export response, the contraction in output is also more front-loaded. These responses contrast with the inflationary price dynamics and gradual export adjustment observed in earlier samples. While the smaller number of major tariff episodes in the postwar period implies a weaker first stage and somewhat less stable estimates, the pattern of responses is nevertheless consistent with a stronger role for demand-side channels in the transmission of tariff shocks in the modern era.[9] In line with this, the estimated output response is also more pronounced in the postwar period.

Importantly, when the postwar period is included as part of a longer sample, the estimated responses largely resemble those obtained from prewar data alone. This reflects the fact that much of the identifying variation in tariff shocks – and many of the largest tariff changes – occur prior to World War II. As a result, longer samples tend to be dominated by prewar variation, with postwar dynamics exerting a more limited influence on the aggregate estimates. Taken together, these findings indicate that while the real effects of tariff shocks are stable and contractionary, the composition of transmission channels has evolved over time, with demand-driven effects likely playing a more prominent role in the postwar period.

5.4. The Role of Domestic and Foreign Policy Responses

What can explain the differences in tariff transmission across samples? A first explanation relates to changes in the monetary regime, in particular the transition from gold-backed monetary systems to fiat money. For much of the pre-World War II period, the United States operated under some form of the gold standard, which fundamentally constrained monetary policy and shaped the adjustment of prices, exchange rates, and external balances following trade shocks. By contrast, in the postwar era nominal exchange rates became increasingly flexible and monetary policy gained greater autonomy, altering the propagation of tariff shocks.

Under the gold standard, nominal exchange rates were effectively fixed, implying

[9] In Appendix E.2.1, we show that the estimated responses in the post-World War II period are more sensitive to the specific events included in the instrument. Using a jackknife exercise that sequentially excludes individual tariff episodes, we find that responses in the full historical sample are stable and not driven by any single event, whereas responses in the post-World War II sample vary more substantially across iterations, reflecting the smaller number of major tariff episodes.

that real exchange rate adjustment could occur only through changes in relative prices. Following a tariff increase, imports fall sharply on impact, improving the trade balance. Because the nominal exchange rate cannot appreciate, this improvement in the external balance leads to gold inflows, an expansion of the domestic money supply, and upward pressure on the price level. As prices increase, U.S. exports become less competitive, leading to a delayed decline in export volumes and a gradual erosion of the initial trade surplus. This adjustment mechanism follows the classical price-specie-flow mechanism described by Hume, whereby trade imbalances under the gold standard are corrected endogenously through gold flows and price adjustments (Hume, 2005).

We confirm these dynamics using a model estimated over the era of gold-backed monetary systems, spanning 1879–1971. Figure 8a presents the corresponding impulse responses for a subset of variables. Consistent with the mechanisms outlined above, a tariff shock generates a gradual real appreciation driven by rising domestic prices, an improvement in the current account, and an expansion in the gold stock reflecting gold inflows. Appendix Figure D.4 shows the full set of impulse responses for this specification.

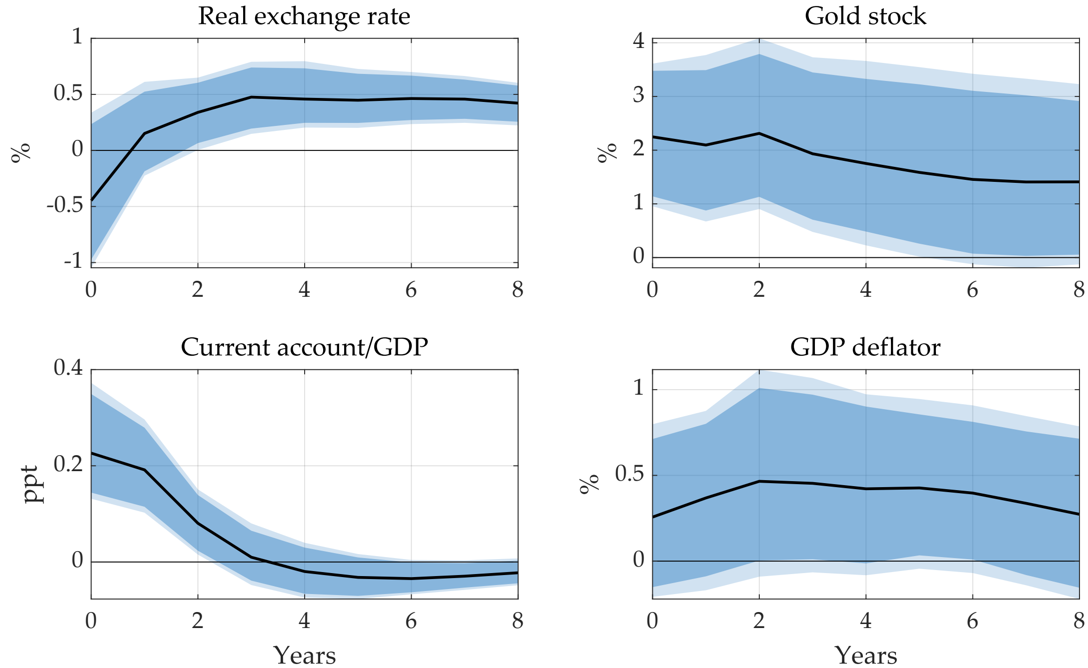
Once the economy operates off the gold standard, the adjustment mechanism changes markedly. With nominal exchange rates free to move, a tariff-induced contraction in import demand can trigger an immediate appreciation of the real exchange rate. This appreciation may offset part of the tariff-driven increase in import prices, dampening pass-through to domestic prices and contributing to a disinflationary price response. Moreover, in the absence of an automatic gold-inflow mechanism, improvements in the trade balance no longer translate into monetary expansion. Instead, the stronger dollar and the associated loss of external competitiveness weigh directly on exports and aggregate demand, placing downward pressure on prices.

We confirm these patterns in Figure 8b, which shows the responses of the real exchange rate and the current account in the pre- and post-World War II periods. In the prewar sample, the real exchange rate appreciates only gradually and the current account improves. In contrast, in the post-World War II period, the real exchange rate appreciates immediately and the current account deteriorates. Differences also arise in the real short rate, which tends to fall in the prewar period but increases post-World War II (see Appendix D.4). This contractionary monetary response may help explain the declines in output and prices observed in the postwar period.

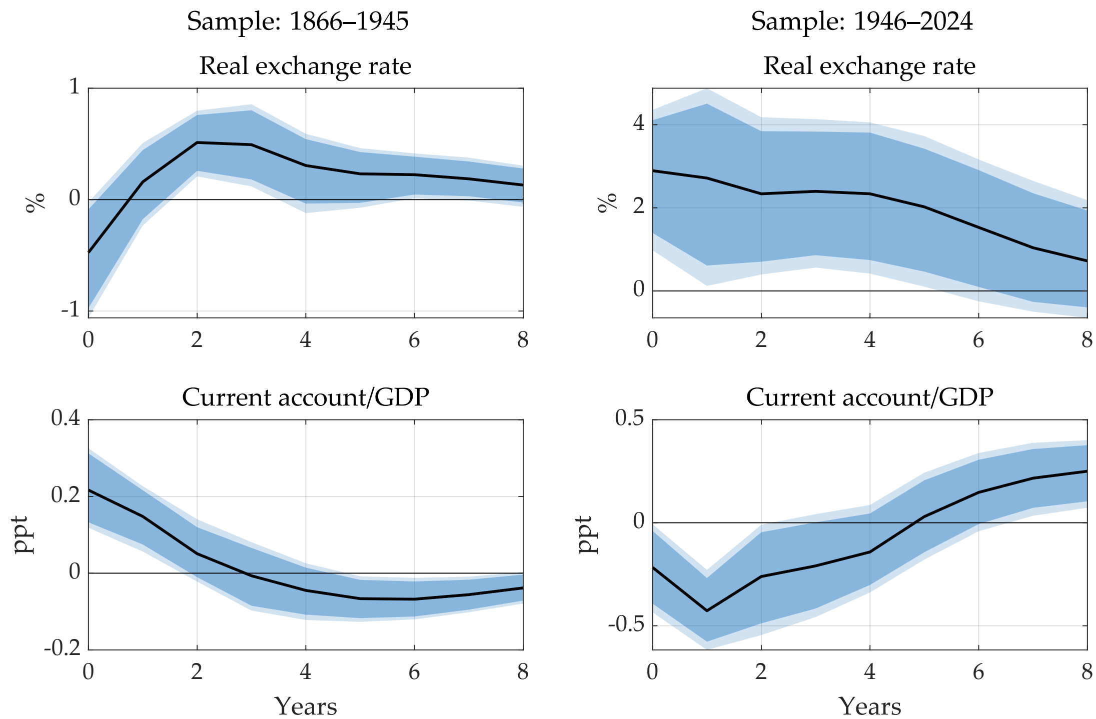
A second, complementary explanation for the changing transmission of tariff shocks lies in the evolution of the international trading system. To assess the role of foreign responses, we construct a measure of foreign tariff rates faced by U.S. exports, defined as an export-weighted average of trading partners' tariffs on U.S. goods (see Appendix B

Figure 8: External Adjustment and Monetary Response

(a) Responses for Era of Gold-Backed Monetary Systems, 1879-1971



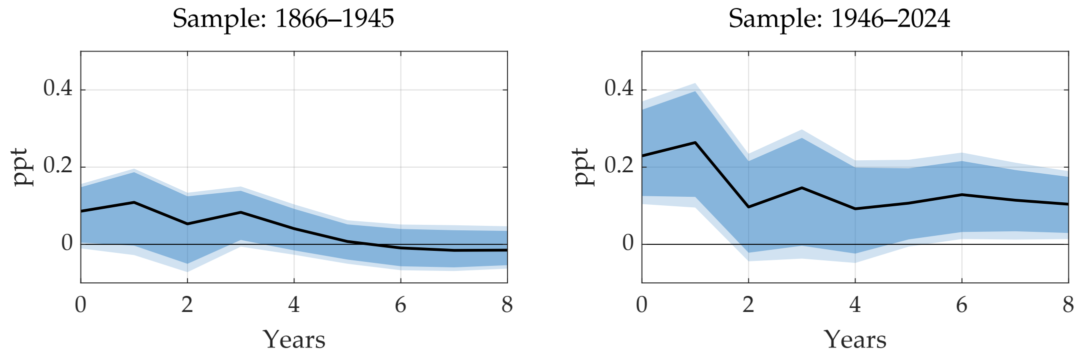
(b) Pre- and Post-World War II



Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact. Panel a: Alternative VAR model (3), including the current account over GDP and the stock of gold in lieu of the import and export variables. Sample: 1879-1971. Panel b: Responses of real exchange rate and current account/GDP obtained using the baseline VAR and the marginal VAR approach, respectively. Sample: 1866–1945 (left panel); 1946–2024 (right panel). Lag order: 2. Solid black line: point estimate. Dark and light shaded areas: 90 and 95 percent confidence bands based on moving-block bootstrap.

for more information). We then compare its response to U.S. tariff shocks in the pre- and post-World War II period.

Figure 9: Responses of Foreign Tariff Rates



Notes: Impulse responses of foreign tariff rates to a U.S. tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact. The foreign tariff rate is defined as an export-weighted average of trading partners' tariffs. Left panel: 1866–1945. Right panel: 1946–2024. Responses are estimated using the VAR model (3), augmented with the foreign tariff rate. Lag order: 2. Solid line: point estimate. Dark and light shaded areas: 90 and 95 percent confidence bands based on moving-block bootstrap.

Figure 9 shows the response of the foreign tariff rate across these two regimes. In the prewar period, the response is muted and not statistically distinguishable from zero. By contrast, in the post-World War II period, the foreign tariffs response increases significantly on impact and remains elevated for several years in response to a U.S. tariff increase. These results are in line with the notion that starting with the RTAA of 1934, tariff policy became increasingly embedded in bilateral and later multilateral trade agreements, with most policy changes taking the form of coordinated tariff reductions. In this period, unilateral tariff increases were more likely to be perceived as deviations from an established cooperative framework in the modern trade regime and therefore to elicit rapid foreign responses via retaliation, reciprocal tariff adjustments, and shifts in sourcing patterns.[10] This stands in contrast to the earlier regime, when tariff changes were more often unilateral and less tightly embedded in a system of reciprocal commitments.[11]

These foreign tariff responses help explain why, in the postwar period, tariffs propagate more strongly through exports and aggregate demand. Rapid foreign retaliation

[10] It is important to note that the post-World War II response in Figure 9 likely represents a lower bound, as the foreign tariff rate is constructed using trading partners' average tariff rates rather than the bilateral tariffs imposed on U.S. goods. As a result, U.S.-specific retaliation is attenuated in this measure. This is especially relevant in the post-World War II period, when country-specific tariffs became more common.

[11] We document a similar pattern when splitting the sample into 1866–1918 and 1919–2024. The difference is more pronounced than in Figure 9 because the pre-World War II period already exhibits a modest retaliatory response, likely reflecting the retaliatory episodes in the 1920s and 1930s, see Appendix D.4.

weakens external demand for U.S. goods, reinforcing the effects of exchange rate appreciation and domestic demand spillovers. This contributes to disinflationary pressures despite higher trade barriers and helps explain the sharper and more immediate export and price responses observed in the modern era. By contrast, in the earlier period, the absence of systematic retaliation and coordination limits the importance of these channels.

Across historical periods, the real effects of tariff shocks are remarkably robust. Tariffs consistently reduce trade volumes and depress economic activity. However, the channels through which tariff shocks propagate seem to have changed over time. Tariffs operate through both supply-side channels, via higher import costs and input prices, and demand-side channels, via exports and aggregate demand. The relative importance of these channels appears to vary with the monetary regime and the structure of the international trading system, shaping in particular the response of prices.

5.5. Sensitivity Analyses

In this section, we perform a comprehensive series of sensitivity checks. Specifically, we assess the importance of the assumptions underlying our VAR approach, including invertibility and the dynamic structure, and investigate the role of potential confounders and anticipation effects.

Relaxing VAR assumptions. Our external-instrument VAR approach relies on a set of potentially restrictive assumptions, most notably invertibility and limited lag-truncation bias. To assess the sensitivity of our results to these assumptions, we re-estimate the effects of tariff shocks using alternative empirical strategies. First, we employ the internal-instrument approach of Plagborg-Møller and Wolf (2019), which relaxes the invertibility requirement while preserving the VAR's dynamic structure. Second, we estimate an LP-IV specification that further relaxes the imposed dynamics.

Figure 10a reports the resulting impulse responses. Across both approaches, the responses are very similar to the baseline responses from the external-instrument VAR. The main difference lies in the precision of the estimates, as reflected in the considerably wider confidence bands. In addition, the terms of trade and the exchange rate response turn out to be slightly attenuated compared to the baseline. Overall, these results suggest that our VAR is rich and flexible enough to recover the shocks of interest and capture U.S. macroeconomic dynamics, while improving precision of the estimated responses.

In the LP-IV framework, we can also assess the sensitivity of the estimates to individual observations and potential pre-trends. A scatter plot of reduced-form outcomes against the instrument reveals no obvious outliers. Moreover, the estimated pre-shock coefficients are small and statistically indistinguishable from zero, providing no evidence of meaningful pre-trends (see Appendix D.5).

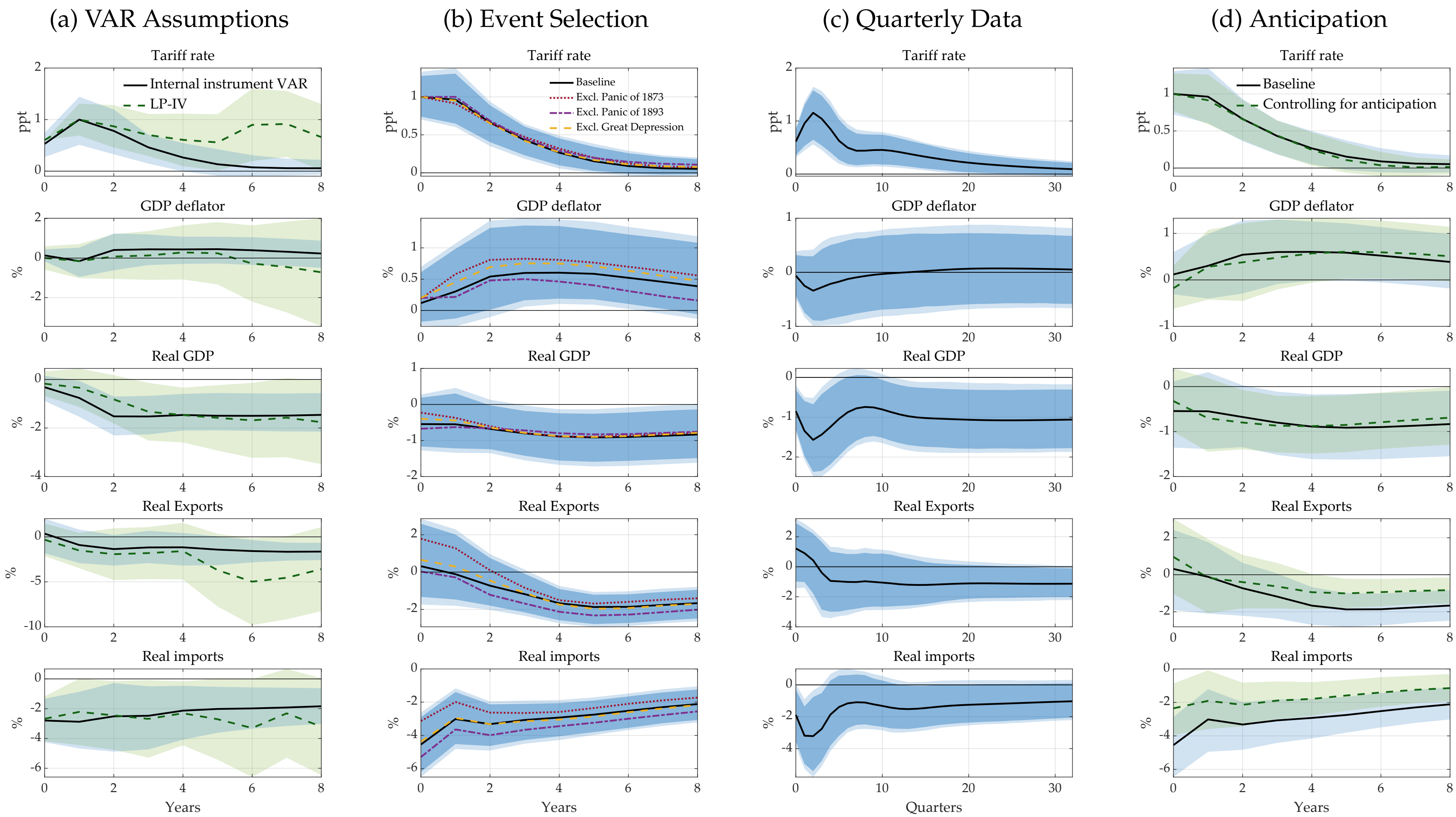
Narrative classification and potential confounders. Our identification strategy relies on a narrative classification of tariff changes designed to isolate policy actions that are plausibly exogenous to contemporaneous macroeconomic conditions. Two related concerns may arise in this context. First, even if a tariff change is plausibly exogenous in terms of its primary motivation, it may still coincide with other macroeconomic shocks – such as financial crises or recessions – that confound the estimated responses. Second, tariff changes may occur alongside broader policy shifts or reflect multiple motivations. Our narrative classification focuses on the primary driver of each reform, but secondary considerations – such as fiscal pressures, partisan policy shifts, or concurrent policy measures – may still introduce residual endogeneity.

To address the first concern, we examine whether our results are driven by tariff changes that occur in periods of major macroeconomic disruptions. As shown in Figure 10b, excluding episodes associated with the Panic of 1873, the Panic of 1893, and the Great Depression yields impulse responses that are closely aligned with the baseline, suggesting that the results are not driven by coincident crisis dynamics. In particular, the results are robust to the exclusion of the Smoot–Hawley Tariff Act, enacted early in the Great Depression.

We further probe this issue using several complementary exercises. First, we construct a quarterly dataset extending back to the mid-nineteenth century. While these series are necessarily noisier, the higher frequency reduces the scope for contemporaneous confounding shocks. Figure 10c shows that the resulting impulse responses closely mirror the baseline estimates. Imports, exports, output, and manufacturing activity all respond in a very similar manner. The main difference is a weaker response of prices; however, this effect is imprecisely estimated and likely reflects increased measurement noise in quarterly historical price data rather than a meaningful change in economic dynamics. Second, a leave-one-out (jackknife) exercise yields responses that remain both qualitatively and quantitatively stable (Appendix E.2), indicating that the results are not driven by a small number of influential episodes. Third, a placebo exercise – assigning tariff shocks to dates without identified policy changes – does not generate systematic movements in trade, output, or prices (Appendix E.4). Taken together, these results suggest that coincident macroeconomic shocks are unlikely to account for our findings.

We address the second concern by examining the role of broader policy confounders.

Figure 10: Sensitivity Analyses



Notes: Sensitivity of impulse responses to a tariff shock. Panel a: Responses based on the internal instrument VAR and LP-IV approach. Solid and dashed lines: point estimate. Shaded areas: 95 percent confidence bands. Panel b: Sensitivity of baseline VAR when excluding potentially problematic events from the instrument. Solid, dashed and dotted lines: point estimate. Dark and light shaded areas: 90 and 95 percent confidence bands based on moving-block bootstrap. Panel c: Responses from quarterly external-instrument VAR model. Sample: 1866Q1-2024Q4. Solid line: point estimate. Dark and light shaded areas: 90 and 95 percent confidence bands based on moving-block bootstrap. Panel d: Sensitivity of baseline VAR when excluding events with long signing-implementation gaps (over three months) from the instrument. Solid and dashed lines: point estimate. Shaded areas: 95 percent confidence bands based on moving-block bootstrap. All responses are normalized to match the peak effect in the baseline, annual external instruments VAR. The full set of responses, including financial variables and manufacturing production, is provided in Appendix E.

We first exclude tariff changes implemented alongside major fiscal measures and control for periods of unified government and partisan control (see Appendix E.2.1 and Den Besten and Känzig, 2026). The estimated responses remain very similar to the baseline, suggesting that coincident policy shifts are unlikely to drive our findings. We also assess the robustness of our narrative classification to alternative treatments of borderline cases. In particular, we consider classifications that vary the handling of episodes in which fiscal imbalances or broader economic pressures may have played a secondary role. As documented in detail in the narrative companion paper (Den Besten and Känzig, 2026), the results are robust to reclassifying these episodes.

Accounting for anticipation effects. A further concern is that some tariff changes may have been anticipated by firms and households prior to their implementation. If economic agents adjust behavior in advance – for instance by front-loading imports or altering production and pricing decisions – standard VAR-based impulse responses centered on the implementation date may understate or mischaracterize the true effects of tariff shocks by missing these preemptive adjustments.

Reassuringly, in many cases, and especially in the earlier part of our sample, the legislative passage and implementation of tariff changes typically occurred within a very narrow time window – limiting the scope for anticipation. Nevertheless, as a robustness check, we re-estimate our baseline specification after excluding tariff events with notable gaps between signing and implementation dates. Figure 10d shows that the resulting impulse responses are very similar to those in the full sample, suggesting that anticipation effects are unlikely to drive our results.

Additional checks. Finally, we conduct a broad set of additional checks to assess the sensitivity of our findings to alternative modeling and measurement choices. These results are reported in Appendix E. We examine the construction of the narrative instrument, including alternative aggregation schemes when mapping events to the annual frequency and specifications that incorporate information on the size of tariff changes rather than only their sign. We also assess the robustness to changes in the VAR specification, varying the set of included variables, the lag order, and the set of deterministic controls. In addition, we verify that our results are not driven by particular data sources by using alternative historical series where available. Across all these exercises, the qualitative responses remain unchanged and the quantitative magnitudes are very similar.

6. Conclusion

This paper studies the macroeconomic effects of tariffs using long-run U.S. historical data. Exploiting the large and frequent tariff changes of the pre-World War II era, we combine a narrative identification strategy with a semi-structural framework to estimate the dynamic causal effects of tariff shocks on output, prices, and trade while addressing the endogeneity of tariff policy.

Tariff increases are contractionary. A one-percentage point increase in the average tariff rate leads to a sizable and persistent decline in real GDP, accompanied by sharp reductions in imports, exports, and manufacturing output. These responses run counter to the protective intent of tariffs and highlight the role of general-equilibrium forces: while imports fall as intended, exports and domestic production decline as well, weakening aggregate economic activity rather than redirecting demand toward domestic producers.

Price responses are muted. Despite higher trade barriers and input costs, inflation reacts only modestly and imprecisely, likely reflecting the offsetting effects of cost pressures and contractionary demand forces operating through reduced trade volumes, exchange-rate appreciation, and weaker external demand.

Accounting for the endogeneity of tariff policy is crucial. Treating all tariff changes as exogenous yields markedly different responses of output, prices, and trade, underscoring the importance of distinguishing policy-driven tariff actions from those taken in response to macroeconomic or fiscal conditions.

Overall, the evidence implies that tariff increases depress economic activity and trade once their indirect, general-equilibrium effects are taken into account. The historical record suggests that the aggregate consequences of tariffs depend not only on their direct impact on import prices, but also on exchange-rate adjustment, foreign responses, and the monetary environment. In this light, the effects of recent tariff hikes are likely to hinge on the strength of these demand-side channels – particularly retaliation and monetary policy – rather than following mechanically from higher trade barriers.

References

Ambrosino, Ludovica, Jenny Chan, and Silvana Tenreyro (2026). "Trade fragmentation, inflationary pressures and monetary policy". *Journal of International Economics*, p. 104219.

Amiti, Mary, Stephen J. Redding, and David E. Weinstein (2019). "The impact of the 2018 tariffs on prices and welfare". *Journal of Economic Perspectives* 33.4, pp. 187–210.

Antolin-Diaz, Juan and Paolo Surico (2025). "The long-run effects of government spending". *American Economic Review* 115.7, pp. 2376–2413.

Auclert, Adrien, Matthew Rognlie, and Ludwig Straub (2025). *The macroeconomics of tariff shocks*. Tech. rep. National Bureau of Economic Research.

Auray, Stéphane, Michael B. Devereux, and Aurélien Eyquem (2025). "Trade wars, nominal rigidities, and monetary policy". *Review of Economic Studies* 92.4, pp. 2228–2270.

Balke, Nathan S. and Robert J. Gordon (1986). *The American Business Cycle: Continuity and Change*. Ed. by Robert J. Gordon. Studies in Business Cycles. Chicago, IL; London, UK: University of Chicago Press.

Bandyopadhyay, Sujan, Domenico Ferraro, and Lorenzo Octavio Vera Bower (2025). "How do tariffs impact the US economy?" *Economics Letters*, p. 112406.

Barnichon, Régis and Aayush Singh (2025). *What Is a Tariff Shock? Insights from 150 years of Tariff Policy*. Tech. rep.

Bergin, Paul R. and Giancarlo Corsetti (2023). "The macroeconomic stabilization of tariff shocks: What is the optimal monetary response?" *Journal of International Economics* 143, p. 103758.

Bianchi, Javier and Louphou Coulibaly (2025). *The optimal monetary policy response to tariffs*. Tech. rep. National Bureau of Economic Research.

Boer, Lukas and Helmut Lütkepohl (2021). "Qualitative versus quantitative external information for proxy vector autoregressive analysis". *Journal of Economic Dynamics and Control* 127, p. 104118.

Boer, Lukas and Malte Rieth (2024). *The macroeconomic consequences of import tariffs and trade policy uncertainty*. IMF Working Papers 2024/013. International Monetary Fund.

Bown, Chad P. (2021). "The US–China trade war and Phase One agreement". *Journal of Policy Modeling* 43.4, pp. 805–843.

Brownlee, W. Elliot (2017). "The Creation of the U.S. Tariff Commission". *A Centennial History of the United States International Trade Commission*. Ed. by Paul R. Bardos. Washington, D.C.: United States International Trade Commission.

Caliendo, Lorenzo, Samuel Kortum, and Fernando Parro (2025). *Tariffs and trade deficits*. Tech. rep. National Bureau of Economic Research.

Carter, Susan B., Scott Sigmund Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch, and Gavin Wright (2006). "International Trade and Exchange Rates". *Historical Statistics of the United States: Earliest Times to the Present, Millennial Edition*. Cambridge: Cambridge University Press. Chap. Ca.

CBO, Congressional Budget Office (1979). *The Effects of the Tokyo Round of Multilateral Trade Negotiations on the US Economy: An Updated View*. US Government Printing Office.

Clemens, Michael A. and Jeffrey G. Williamson (2004). "Why did the tariff–growth correlation change after 1950?" *Journal of Economic Growth* 9.1, pp. 5–46.

Cloyne, James (2013). "Discretionary tax changes and the macroeconomy: new narrative evidence from the United Kingdom". *American Economic Review* 103.4, pp. 1507–28.

Cox, Lydia (2025). "The long-term impact of steel tariffs on US manufacturing".

Den Besten, Tamar and Diego R. Känzig (2025). "The Macroeconomic Effects of Tariffs: Evidence From US Historical Data". *Available at SSRN 5938997*.

— (2026). "Tariff Changes in the U.S. from 1840–2024: A Narrative Analysis".

Dobson, John M. (1976). *Two Centuries of Tariffs: The Background and Emergence of the U.S. International Trade Commission*. USITC Publication. United States International Trade Commission.

Dolan, Lindsay R., Robert M. Kubinec, Daniel L. Nielson, and Jiakun Jack Zhang (2025). "Tariffs and corporate political activity: a survey experiment on US businesses". *Business and Politics*, pp. 1–22.

Eaton, Jonathan and Samuel Kortum (2002). "Technology, geography, and trade". *Econometrica* 70.5, pp. 1741–1779.

Eichengreen, Barry (1999). "The Political Economy of the Smoot-Hawley Tariff". *International Political Economy*. Taylor & Francis Ltd. Chap. 2, pp. 37–46.

Fabricant, Solomon (1940). *The Output of Manufacturing Industries, 1899–1937*. New York: National Bureau of Economic Research.

Fajgelbaum, Pablo D., Pinelopi K. Goldberg, Patrick J. Kennedy, and Amit K. Khandelwal (2020). "The return to protectionism". *The Quarterly Journal of Economics* 135.1, pp. 1–55.

Fieldhouse, Andrew J. and Karel Mertens (2023). "The returns to government R&D: Evidence from US appropriations shocks".

Fisk, George M. (1910). "The Payne-Aldrich Tariff". *Political Science Quarterly* 25.1, pp. 35–68.

Flaaen, Aaron and Justin Pierce (2024). "Disentangling the Effects of the 2018–2019 Tariffs on a Globally Connected U.S. Manufacturing Sector". *Review of Economics and Statistics*.

Flaherty, Jane (2001). "Incidental Protection: An Investigation of the Morrill Tariff". *Essays in Economic & Business History* 19, pp. 103–118.

Fornaro, Luca and Martin Wolf (2025). "Tariffs and technological hegemony".

Franconi, Alessandro and Lukas Hack (2025). *Import Tariffs and the Systematic Response of Monetary Policy*. SSRN working paper.

Frickey, Edwin (1947). *Production in the United States, 1860–1914*. Harvard Economic Studies. Cambridge, MA: Harvard University Press, p. 54.

Furceri, Davide, Swarnali A. Hannan, Jonathan D. Ostry, and Andrew K. Rose (2022). "The macroeconomy after tariffs". *The World Bank Economic Review* 36.2, pp. 361–381.

GAO, United States General Accounting Office (1994). *The General Agreement on Tariffs and Trade: Uruguay Round Final Act Should Produce Overall U.S. Economic Gains*. Tech. rep. Washington, D.C.: U.S. General Accounting Office.

Gertler, Mark and Peter Karadi (2015). "Monetary policy surprises, credit costs, and economic activity". *American Economic Journal: Macroeconomics* 7.1, pp. 44–76.

Greenland, Andrew and John Lopresti (2024). "Trade policy as an exogenous shock: Focusing on the specifics".

Hume, David (2005). "On the balance of trade". *Gold standard in theory & history*. Routledge, pp. 31–37.

Irwin, Douglas A. (1995). "The GATT in Historical Perspective". *The American Economic Review* 85.2, pp. 323–328.

— (1998a). "Changes in US tariffs: the role of import prices and commercial policies". *The American Economic Review* 88.4, pp. 1015–1026.

— (1998b). "Higher Tariffs, Lower Revenues? Analyzing the Fiscal Aspects of "The Great Tariff Debate of 1888"". *The Journal of Economic History* 58.1, pp. 59–72.

— (2000). "Did late-nineteenth-century US tariffs promote infant industries? Evidence from the tinplate industry". *The Journal of Economic History* 60.2, pp. 335–360.

— (2013). "The Nixon shock after forty years: the import surcharge revisited". *World Trade Review* 12.1, pp. 29–56.

— (2017a). *Clashing over commerce: A history of U.S. trade policy*. University of Chicago Press.

— (2017b). "The False Promise of Protectionism: Why Trump's Trade Policy Could Backfire". *Foreign Affairs* 96.3, pp. 45–56.

— (2020). "Trade policy in American economic history". *Annual Review of Economics* 12.1, pp. 23–44.

Irwin, Douglas A. and Randall S. Kroszner (1996). "Log-rolling and economic interests in the passage of the Smoot-Hawley Tariff". *Carnegie-Rochester Conference Series on Public Policy*. Vol. 45. Elsevier, pp. 173–200.

Itskhoki, Oleg and Dmitry Mukhin (2025). *The optimal macro tariff*. Tech. rep. National Bureau of Economic Research.

Jentsch, Carsten and Kurt G. Lunsford (2019). "The dynamic effects of personal and corporate income tax changes in the United States: Comment". *American Economic Review* 109.7, pp. 2655–78.

Johnson, Harry G. (1953). "Optimum tariffs and retaliation". *The Review of Economic Studies* 21.2, pp. 142–153.

Jordà, Òscar, Moritz Schularick, and Alan M. Taylor (2013). "When credit bites back". *Journal of Money, Credit and Banking* 45.s2, pp. 3–28.

— (2015). "Betting the house". *Journal of International Economics* 96, S2–S18.

— (2016). "The great mortgaging: housing finance, crises and business cycles". *Economic Policy* 31.85, pp. 107–152.

Jordà, Òscar, Sanjay R. Singh, and Alan M. Taylor (2024). "The long-run effects of monetary policy". *Review of Economics and Statistics*, pp. 1–49.

Klein, Alexander and Christopher M. Meissner (2024). *Did tariffs make American manufacturing great? New evidence from the Gilded Age*. Tech. rep. National Bureau of Economic Research.

Lerdau, Enrique (1957). "On the Measurement of Tariffs: The U.S. Over Forty Years". *Economia Internazionale* 10, pp. 232–244.

Lipsey, Robert E. (1963). *Price and Quantity Trends in the Foreign Trade of the United States*. Princeton, NJ: Princeton University Press.

McKinley, William (1890). *Congressional Record—House: Tariff Bill*. U.S. Government Printing Office.

Mertens, Karel and Morten O. Ravn (2013). "The dynamic effects of personal and corporate income tax changes in the United States". *American Economic Review* 103.4, pp. 1212–47.

Miranda-Agrippino, Silvia and Giovanni Ricco (2018). "Identification with external instruments in structural VARs under partial invertibility".

Monacelli, Tommaso (2025). *Tariffs and Monetary Policy*. Tech. rep. CEPR Discussion Paper.

Montiel Olea, José Luis and Carolin Pflueger (2013). "A robust test for weak instruments". *Journal of Business & Economic Statistics* 31.3, pp. 358–369.

Montiel Olea, José Luis and Mikkel Plagborg-Møller (2020). "Local Projection Inference is Simpler and More Robust Than You Think".

Müller, Karsten, Chenzi Xu, Mohamed Lehbib, and Ziliang Chen (2025). *The Global Macro Database: A New International Macroeconomic Dataset*. Working Paper 33714. National Bureau of Economic Research.

North, Douglass C. (1961). *The Economic Growth of the United States, 1790–1860*. Englewood Cliffs, NJ: Prentice Hall.

O'Rourke, Kevin H. (2000). "Tariffs and growth in the late 19th century". *The Economic Journal* 110.463, pp. 456–483.

Officer, Lawrence H. and Samuel H. Williamson (2025). *Annual Wages in the United States, 1774–Present*. MeasuringWorth.

Ossa, Ralph (2014). "Trade wars and trade talks with data". *American Economic Review* 104.12, pp. 4104–4146.

Pastor, Robert A. (1980). *Congress and the Politics of U.S. Foreign Economic Policy, 1929–1976*. Berkeley: University of California Press.

Plagborg-Møller, Mikkel (2022). "Discussion of "Narrative Restrictions and Proxies" by Raffaella Giacomini, Toru Kitagawa, and Matthew Read". *Journal of Business & Economic Statistics* 40.4, pp. 1434–1437.

Plagborg-Møller, Mikkel and Christian K. Wolf (2019). "Local projections and VARs estimate the same impulse responses".

Ramey, Valerie A. (2011). "Identifying government spending shocks: It's all in the timing". *The Quarterly Journal of Economics* 126.1, pp. 1–50.

— (2016). "Macroeconomic shocks and their propagation". *Handbook of Macroeconomics* 2, pp. 71–162.

Ramey, Valerie A. and Matthew D. Shapiro (1998). "Costly capital reallocation and the effects of government spending". *Carnegie-Rochester Conference Series on Public Policy*. Vol. 48. Elsevier, pp. 145–194.

Ramey, Valerie A. and Sarah Zubairy (2018). "Government spending multipliers in good times and in bad: evidence from US historical data". *Journal of Political Economy* 126.2, pp. 850–901.

Rodríguez-Clare, Andrés, Mauricio Ulate, and Jose P. Vasquez (2025). *The 2025 trade war: Dynamic impacts across US States and the global economy*. Tech. rep. National Bureau of Economic Research.

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, pp. 121–170.

Romer, Christina D. and David H. Romer (2004). "A new measure of monetary shocks: Derivation and implications". *American Economic Review* 94.4, pp. 1055–1084.

— (2010). "The macroeconomic effects of tax changes: estimates based on a new measure of fiscal shocks". *American Economic Review* 100.3, pp. 763–801.

— (2023). "Presidential address: Does monetary policy matter? The narrative approach after 35 years". *American Economic Review* 113.6, pp. 1395–1423.

Schmitt-Grohé, Stephanie and Martín Uribe (2025). *Transitory and permanent import tariff shocks in the United States: An empirical investigation*. Tech. rep. National Bureau of Economic Research.

Schularick, Moritz and Solomos Solomou (2011). "Tariffs and economic growth in the first era of globalization". *Journal of Economic Growth* 16.1, pp. 33–70.

Simon, Matthew (1960). "The United States balance of payments, 1861-1900". *Trends in the American economy in the nineteenth century*, pp. 629–716.

Steinberg, Joseph B. (2025). *Tariffs, Manufacturing Employment, and Supply Chains*. Tech. rep. National Bureau of Economic Research.

Stock, James H. and Mark W. Watson (2018). "Identification and estimation of dynamic causal effects in macroeconomics using external instruments". *The Economic Journal* 128.610, pp. 917–948.

Tarbell, Ida Minerva (1911). *The Tariff in Our Times*. Macmillan.

Taussig, Frank W. (1910). *The Tariff History of the United States*. 5th. New York and London: G.P. Putnam's Sons / The Knickerbocker Press.

U.S. Census Bureau (1949). *Historical Statistics of the United States, 1789–1945: A Supplement to the Statistical Abstract of the United States*. Washington, D.C.: U.S. Government Printing Office.

Werning, Iván, Guido Lorenzoni, and Veronica Guerrieri (2025). *Tariffs as Cost-Push Shocks: Implications for Optimal Monetary Policy*. Tech. rep. National Bureau of Economic Research.

Online Appendix

The Macroeconomic Effects of Tariffs

Tamar den Besten
Régis Barnichon
Diego R. Känzig
Aayush Singh

Contents

A. Narrative Analysis

We present an overview of our narrative analysis in Table A.1. The majority of the analysis is based on the comprehensive account of tariff changes and their broader political and economic context provided by Irwin (2017) and Taussig (1910) for earlier tariff changes, supplemented with a wide range of additional historical and legislative sources. The table provides a concise summary of the events included in our analysis. The full historical context and the detailed reasoning underlying the classification of tariff changes as endogenous or exogenous are documented in the companion paper (Den Besten and Känzig, 2026).

Table A.1: Summary of Narrative Analysis

	Tariff	Implementation date	Direction	Main motivation	Classification
1	Tariff of 1842	August 30, 1842	Increase	Fiscal imbalances	Endogenous
2	Walker Tariff	December 1, 1846	Decrease	Trade ideology; Distributional	Exogenous
3	Tariff of 1857	July 1, 1857	Decrease	Fiscal imbalances	Endogenous
4	Morrill Tariff	April 1, 1861	Increase	Fiscal imbalances	Endogenous
5	Revenue Act of 1861	August 5, 1861	Increase	Spending-driven	Endogenous
6	Act of December 24, 1861	December 24, 1861	Increase	Spending-driven	Endogenous
7	Tariff Act of July 14, 1862	August 1, 1862	Increase	Spending-driven	Endogenous
8	Resolution of April 29, 1864	April 29, 1864	Increase	Spending-driven	Endogenous
9	Tariff Act of 1864	July 1, 1864	Increase	Spending-driven	Endogenous
10	Act of March 3, 1865	April 1, 1865	Increase	Spending-driven	Endogenous
11	Tariff of 1870	January 1, 1871	Decrease	Political	Exogenous
12	Tariff of 1872	August 1, 1872	Decrease	Political	Exogenous
13	Tariff of 1875	March 3, 1875	Increase	Political	Exogenous
14	Mongrel Tariff	July 1, 1883	Decrease	Distributional; Political	Exogenous
15	McKinley Tariff	October 6, 1890	Increase	Trade ideology	Exogenous
16	Wilson–Gorman Tariff	August 1, 1894	Decrease	Trade ideology; Distributional	Exogenous
17	Dingley Tariff	July 24, 1897	Increase	Fiscal imbalances; External imbalances	Endogenous
18	Payne–Aldrich Tariff	August 6, 1909	Decrease	Trade ideology	Exogenous
19	Underwood-Simmons Tariff Act	October 4, 1913	Decrease	Trade ideology; Distributional	Exogenous
20	Emergency Tariff Act	May 28, 1921	Increase	Countercyclical	Endogenous
21	Fordney–McCumber Tariff Act	September 22, 1922	Increase	Trade ideology	Exogenous[1]
22	Smoot-Hawley Tariff Act	June 18, 1930	Increase	Distributional	Exogenous
23	Reciprocal Tariff Act	September 3, 1934[2]	Decrease	Countercyclical	Endogenous
24	GATT: Geneva Round I	January 1, 1948	Decrease	Trade ideology; Political	Exogenous
25	GATT: Annecy Round	January 1, 1950	Decrease	Trade ideology; Political	Exogenous
26	GATT: Torquay Round	June 6, 1951	Decrease	Trade ideology; Political	Exogenous
27	GATT: Geneva II Round	June 30, 1956	Decrease	Trade ideology; Political	Exogenous

	Tariff	Implementation date	Direction	Main motivation	Classification
28	GATT: Dillon Round	December 31, 1962	Decrease	Trade ideology; Political	Exogenous
29	GATT: Kennedy Round	January 1, 1968	Decrease	Trade ideology; Political	Exogenous
30	Nixon Surcharge	August 16, 1971	Increase	External imbalances	Endogenous
31	Ford's Tax on Imported Oil[3]	February 1, 1975	Increase	External imbalances	Endogenous
32	GATT: Tokyo Round	January 1, 1980	Decrease	Trade ideology; Political	Exogenous
33	GATT: Uruguay Round	January 1, 1995	Decrease	Trade ideology; Political	Exogenous
34	2018 Trump Tariffs	February 7, 2018[4]	Increase	Trade ideology; Political	Exogenous
35	2019 Trump Tariffs	May 10, 2019[5]	Increase	Trade ideology; Political	Exogenous

[1] We exclude this event from the exogenous series in our annual specification, because its timing coincides with the period covered by the distributed proxy for the endogenous Emergency Tariff.

[2] The Reciprocal Tariff Act was signed on June 12, 1934, and the first trade agreement under the act (between the United States and Cuba) went into effect on September 3, 1934.

[3] Officially an *import fee*, but it is included here because it functioned like a tariff.

[4-5] These dates correspond to the first tariff action implemented in each year.

B. Data

In this appendix, we provide details on the construction of our annual and quarterly datasets. We also present descriptive evidence on U.S. exposure to international trade, as well as the composition of imports and exports over time.

B.1. Annual Data

Table B.1 summarizes the sources used to construct the annual dataset. We build the historical series by starting from the most recent source and splicing it to earlier data at the first year of overlap. We apply a growth-rate splice to preserve the level of the recent data while extending the series backward using the growth rates from the historical sources.

Table B.1: Description and Sources of Annual Data

Variable	Description	Source	Sample
Baseline variables			
Tariff rate	Average tariff rate on dutiable imports, calculated as: duties calculated/value of dutiable imports	U.S. Census Bureau (1949) Carter et al. (2006) United States International Trade Commission	1840–1945 1946–1988 1989–2024
Terms of trade	Terms of Trade, calculated as: export price index/import price index	North (1961) Simon (1960) Lipsey (1963) U.S. Bureau of Economic Analysis	1840–1860 1860–1879 1879–1929 1929–2024
Real exports	Real exports of goods (merchandise), deflated by the GDP deflator	U.S. Census Bureau (1949) Carter et al. (2006) United States International Trade Commission	1840–1945 1945–1989 1989–2024
Real imports	Real imports of goods (merchandise), deflated by the GDP deflator	U.S. Census Bureau (1949) Carter et al. (2006) United States International Trade Commission	1840–1945 1945–1989 1989–2024
Real GDP	Real gross domestic product	Müller et al. (2025)	1840–2024
Manufacturing production	Manufacturing production index	Davis (2004) Frickey (1947) Fabricant (1940) Board of Governors of the Federal Reserve System	1840–1860 1860–1899 1899–1919 1919–2024
Real compensation	Real hourly compensation of manufacturing production workers, deflated by the GDP deflator	Officer and Williamson (2025)	1840–2024
GDP deflator	Gross domestic product: implicit price deflator	Müller et al. (2025)	1840–2024

Variable	Description	Source	Sample
Real exchange rate	Real effective exchange rate of the U.S. Dollar against a narrow basket of trading partners	FX & CPI: Müller et al. (2025); Import shares[1] Carter et al. (2006) Bank for International Settlements	1850–1964 1964–2024
Interest rate spread	Interest rate spread, calculated as the difference between long- and short-term interest rates	Müller et al. (2025)	1840–2024
Additional variables			
Real government debt	Real general government debt, deflated by GDP deflator	Müller et al. (2025)	1840–2024
Real government revenue	Federal government total receipts, deflated by GDP deflator	Müller et al. (2025) NBER Macrohistory Database U.S. Bureau of Economic Analysis	1840–1879 1879–1929 1929–2024
Real government expenditure	Real general government expenditure, deflated by GDP deflator	Müller et al. (2025)	1840–2024
SP500	Historical composite price index of the U.S. stock market	Schwert (1990) Global Financial Data (Finaeon)	1850–1871 1871–2024
Real energy commodity price index	Global real prices of natural gas, coal, and petroleum, weighted by US consumption shares	Real prices: Jacks (2019); consumption shares: Wright (2006)	1850-2024
Producer price index	Producer price index, all commodities (extended using historical wholesale price indexes)	Warren and Pearson (1932) Hanes (1998) U.S. Bureau of Labor Statistics	1850–1860 1860–1913 1913–2024
Consumer price index	Consumer price index	Müller et al. (2025)	1840–2024
Real short rate	Real short-term interest rate, calculated as the short-term nominal interest rate minus the inflation rate	Müller et al. (2025)	1840–2024
Current account/GDP	Current account balance as a share of GDP	Müller et al. (2025)	1840–2024
Gold stock	U.S. monetary gold stock, end of year, converted to fine ounces	Bordo (2006)	1879–1971
Foreign tariff rate	Export-weighted average tariff rate on total imports of thirteen major U.S. trading partners[2]	*See sources for tariff rates and export shares below*	
Tariff rates	Customs revenue as a share of total imports of goods	Imlah (1958), Lévy-Leboyer and Bourguignon (1985), Mitchell (2013), Fouquin and Hugot (2016), OECD Global Revenue Statistics and International Merchandise Trade Statistics	1840–2024[3]

Export shares	Export share of country c in total U.S. exports	Carter et al. (2006), Fouquin and Hugot (2016)	1840–1988
		United States International Trade Commission	1989–2024

[1] Unlike the BIS methodology, which uses both export and import shares to determine relative country weights (Klau and Fung, 2006), we rely only on import shares due to limited availability of historical export data. The resulting series is nevertheless very similar to the BIS USD REER over the overlapping period.

[2] These include the United Kingdom, France, Germany, Italy, Canada, Switzerland, Belgium, the Netherlands, Portugal, Spain, Japan, Mexico, and Brazil, where data are available.

[3] The choice of data source varies across countries and years.

B.2. Quarterly Data

As discussed in the main text, we also construct a quarterly dataset to improve the timing of tariff shocks and reduce concerns that annual tariff changes may coincide with other contemporaneous shocks. Because higher-frequency data are limited in earlier periods, we rely on a combination of (i) manual digitization of historical sources (including microfiche from the Northwestern Library and scanned PDFs of official reports), and (ii) temporal disaggregation of annual series using higher-frequency indicators.

Regarding the latter, we convert low-frequency economic variables (e.g., annual GDP) into high-frequency measures (e.g., quarterly GDP) using related high-frequency indicators (e.g., industrial production). The approach ensures that short-run fluctuations in the indicator anchor the quarterly path of the target series between annual observations. We use high-frequency indicators that are most relevant for the variable we aim to temporally disaggregate, as in Antolin-Diaz and Surico (2025). Following Balke and Gordon (1986), we interpolate using the Chow–Lin method (Chow and Lin, 1971), which regresses the low-frequency variable on the indicator and models the residuals as an AR(1) process. The method uses generalized least squares (GLS) to account for autocorrelation and ensures that the resulting quarterly estimates preserve observed annual totals exactly. We implement the method using the package by Jaramillo-Vera (2025), and use the *chow-lin-opt* variant, which estimates the AR(1) parameter via maximum likelihood.

Where appropriate we seasonally adjust quarterly data using the U.S. Census Bureau's X-13ARIMA-SEATS method. Table B.2 presents the quarterly sources used to construct our dataset.

Table B.2: Description and Sources of Quarterly Data

Variable	Description	Source	Sample
Tariff rate	Average tariff rate on dutiable imports, calculated as: duties calculated/value of dutiable imports		
Dutiable imports	Value of dutiable merchandise entered for consumption from foreign ports	Report of the Secretary of the Treasury on the State of the Finances	1855Q3–1866Q2

Variable	Description	Source	Sample
	Total value of imports of merchandise: Dutiable	Monthly Summary of the Foreign Commerce of the United States	1866Q3–1961Q2
	Annual dutiable imports for consumption, interpolated using quarterly imports for consumption	United States Imports of Merchandise for Consumption	1961Q2–1966Q4
		U.S. Imports	1967Q1
	Imports for consumption: dutiable	Highlights of U.S. Export and Import Trade. FT990[1]	1967Q1–1988Q4
	Imports for consumption: dutiable value	United States International Trade Commission	1989Q1–2024Q4
Duties collected[2]	Receipts from customs into the treasury	Report of the Secretary of the Treasury on the State of the Finances	1850Q1–1871Q3
	Federal budget receipts, customs for the United States [M15001USM144NNBR]	NBER Macrohistory Database	1871Q3–1959Q1
	Gross receipts: customs	Monthly Treasury Statement	1959Q1–2024Q4
Terms of trade	Terms of Trade, calculated as: export price index/import price index		
Export price index	Annual export price index, interpolated using the U.S. wholesale price index [M0448AUSM323NNBR]	NBER Macrohistory Database	1850Q1–1880Q1
	Annual export price index, interpolated using the import price index for Great Britain [M04102GBM318NNBR]	NBER Macrohistory Database	1880Q1–1913Q4
	Annual export price index, interpolated using the U.S. wholesale price index [M0448AUSM323NNBR]	NBER Macrohistory Database	1913Q4–1919Q1
	Price deflator of exports	Balke and Gordon (1986)	1919Q1–1933Q3
	Price index of exports of domestic merchandise [M04196USM345NNBR]	NBER Macrohistory Database	1933Q3–1947Q1
	Exports of goods (chain-type price index) [B253RG3Q086SBEA]	U.S. Bureau of Economic Analysis	1947Q1–2024Q4
Import price index	Annual import price index, interpolated using the U.S. wholesale price index [M0448AUSM323NNBR]	NBER Macrohistory Database	1850Q1–1880Q1
	Annual import price index, interpolated using the export price index for Great Britain [M04109GBM318NNBR]	NBER Macrohistory Database	1880Q1–1913Q4
	Annual import price index, interpolated using the U.S. wholesale price index [M0448AUSM323NNBR]	NBER Macrohistory Database	1913Q4–1919Q1
	Price deflator of imports	Balke and Gordon (1986)	1919Q1–1933Q3
	Price index of imports for consumption [M04197USM345NNBR]	NBER Macrohistory Database	1933Q3–1947Q1
	Imports of goods (chain-type price index) [B255RG3Q086SBEA]	U.S. Bureau of Economic Analysis	1947Q1 - 2024Q4

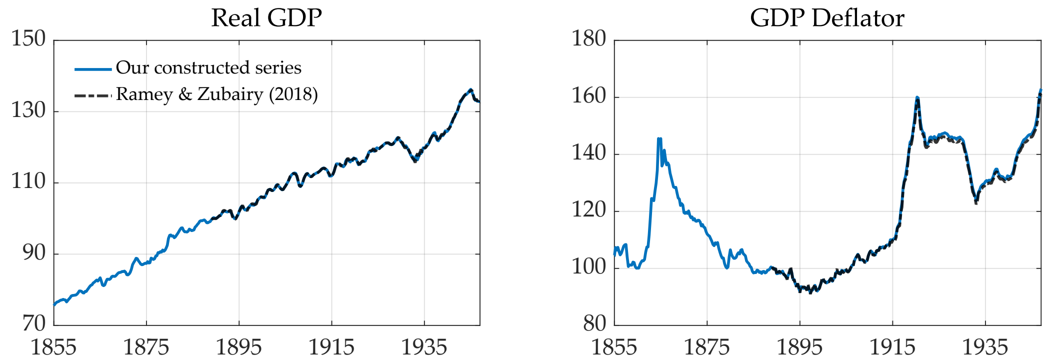
Variable	Description	Source	Sample
Real exports	Real exports of goods (merchandise), deflated by the GDP deflator		
	Annual exports of goods interpolated using exports from New York to foreign ports, exclusive of specie	The Merchants' Magazine and Commercial Review	1851Q1–1866Q3
	Total exports for the United States [(M07023USM144NNBR]	NBER Macrohistory Database	1866Q3–1955Q1
	Total exports, goods	OECD, International Merchandise Trade Statistics (IMTS)	1955Q1–2024Q4
Real imports	Real imports of goods (merchandise), deflated by the GDP deflator		
	Total imports of merchandise	Report of the Secretary of the Treasury on the State of the Finances	1855Q3–1866Q2
	Total imports for the United States [M07028USM144NNBR]	NBER Macrohistory Database	1866Q3–1955Q1
	Total imports, goods	OECD, International Merchandise Trade Statistics (IMTS)	1955Q1–2024Q4
Real GDP	Real gross domestic product		
	Annual GDP, interpolated using index of American business activity [M12003USM516NNBR][3]	NBER Macrohistory Database	1855Q1–1875Q1
	Annual GDP, interpolated using index of industrial production and trade (W. Persons) [M1204AUSM516SNBR]	NBER Macrohistory Database	1875Q1–1884Q3
	Annual GDP, interpolated using index of industrial production	Miron and Romer (1990)	1884Q3–1919Q1
	Annual GDP, interpolated using industrial production: total index	Board of Governors of the Federal Reserve System	1919Q1–1947Q1
	Real gross domestic product [GDPC1]	U.S. Bureau of Economic Analysis	1947Q1–2024Q4
Manufacturing production	Manufacturing production index		
	Annual manufacturing production index, interpolated using index of American business activity [M12003USM516NNBR]	NBER Macrohistory Database	1855Q1–1875Q1
	Annual manufacturing production index, interpolated using index of industrial production and trade (W. Persons) [M1204AUSM516SNBR]	NBER Macrohistory Database	1875Q1–1884Q3
	Annual manufacturing production index, interpolated using index of industrial production	Miron and Romer (1990)	1884Q3–1919Q1
	Industrial production: manufacturing (SIC) [IPB00004SQ]	Board of Governors of the Federal Reserve System	1919Q1–2024Q4

Variable	Description	Source	Sample
Real compensation	Real hourly compensation of manufacturing production workers, deflated by the GDP deflator		
	Annual nominal hourly compensation of manufacturing production workers, interpolated using nominal GDP [4]	See the sources for quarterly real GDP and the GDP deflator	1855Q1-1914Q1
	Annual nominal hourly compensation of manufacturing production workers, interpolated using average weekly earnings, manufacturing [M08261USM052NNBR]	NBER Macrohistory Database	1914Q1–1939Q1
	Annual nominal hourly compensation of manufacturing production workers, interpolated using average hourly earnings of production and nonsupervisory employees, manufacturing [AHEMAN]	U.S. Bureau of Labor Statistics	1939Q1–2024Q4
GDP deflator	Gross domestic product: implicit price deflator		
	Annual GDP deflator, interpolated using the U.S. wholesale price index	NBER Macrohistory Database Balke and Gordon (1986)	1850Q1–1875Q1 1875Q1–1947Q1
	Gross domestic product: implicit price deflator [GDPDEF]	U.S. Bureau of Economic Analysis	1947Q1–2024Q4
Real exchange rate	Real effective exchange rate of the U.S. Dollar against a narrow basket of trading partners		
	Annual real effective exchange rate, interpolated using the real GBP/USD exchange rate	Bank for International Settlements	1850Q1–1964Q1
	Real effective exchange rate, United States, narrow basket	Bank for International Settlements	1964Q1–2024Q4
Interest rate spread	Interest rate spread, calculated as the difference between long- and short-term interest rates		
Short-term rate	AA 3 month Non-Financial Commercial Paper	Global Financial Data (Finaeon)	1850Q1–1930Q4
	90-day Secondary Market Treasury Bills	Global Financial Data (Finaeon)	1930Q4–2024Q4
Long-term rate	10-year bond constant maturity yield	Global Financial Data (Finaeon)	1850Q1–2024Q4

[1] Data on dutiable imports are manually digitized from the historical reports referenced above. For issues of the *Highlights of U.S. Export and Import Trade* not available in PDF form, we rely on microfiche supplied by the Northwestern University Library.
[2] In the annual specification, the tariff-rate numerator is duties *calculated*, rather than duties *collected*. To ensure comparability, we interpolate annual duties calculated using the quarterly duties collected.
[3] The Index of American Business Activity and Persons' Index of Industrial Production and Trade are reported as *percent of trend*. Following Balke and Gordon (1986), we first detrend the annual series using benchmark years in which the high-frequency index is close to normal. We then apply the Chow–Lin procedure and finally reintroduce the secular trend using the same benchmarks.
[4] We construct quarterly nominal GDP using the quarterly real GDP and GDP deflator series described above.

Ramey and Zubairy (2018) construct quarterly real GDP and GDP deflator series for the United States covering 1889–2015. Because our sample extends further back in time, we construct our own series to ensure that the methodology is consistent over the full historical period. Despite some differences in the construction of the series[1], Figure B.1 shows that our results are comparable and consistent with the off-the-shelf data provided by Ramey and Zubairy (2018).

Figure B.1: Comparison of Our Data vs. Off-the-Shelf Data



Notes: All series are expressed in logs and rebased to 1889 to match the start of the off-the-shelf data. Sample: 1855–1947. Solid blue lines show our constructed series (see Table B.2). Dashed black lines show the historical series provided by Ramey and Zubairy (2018).

B.3. Descriptive Statistics and Figures

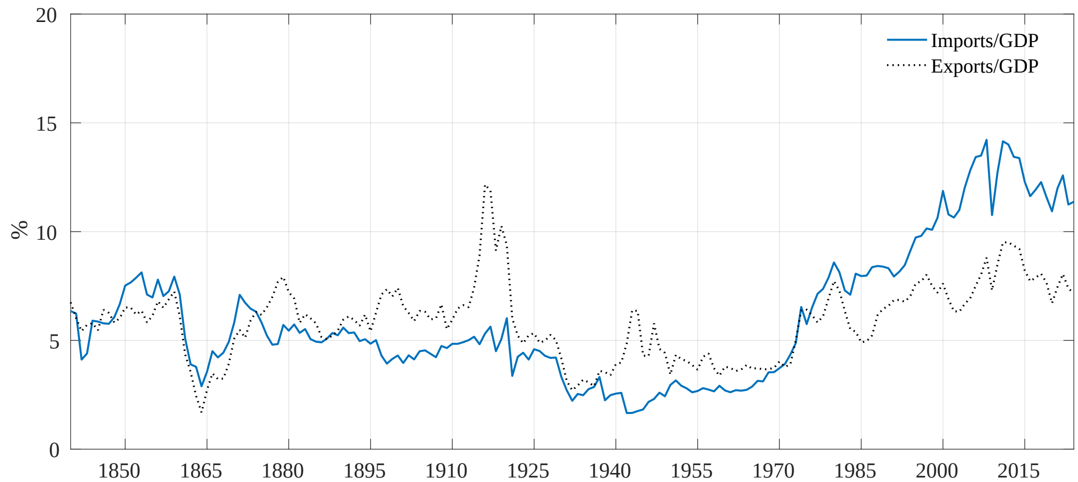
This appendix provides additional descriptive evidence on the evolution of U.S. trade exposure and the composition of imports and exports.

Figure B.2 shows that the United States has historically been meaningfully exposed to international trade. Imports and exports of goods typically account for between 5 and 10 percent of GDP over much of the sample. Trade openness increased markedly in the post–World War II period, with imports rising more strongly than exports. By the early 2000s, goods imports reached nearly 15 percent of GDP, while exports approached 10 percent. This increase in import penetration suggests that the macroeconomic effects of tariff changes may have become more pronounced over time, as a larger share of domestic expenditure is directly affected by trade barriers.

Figures B.3 and B.4 document the composition of imports and exports by broad end-use category since 1925. Prior to World War II, the composition of trade is relatively stable. Imports are dominated by industrial supplies and materials, while food products account

[1] For example, Ramey and Zubairy (2018) use the proportional Denton procedure to interpolate existing annual series while we use Chow–Lin; Ramey and Zubairy (2018) use the quarterly GNP deflator constructed by Balke and Gordon (1986), which is an interpolated series using U.S. wholesale prices, while we interpolate the GDP deflator directly using the U.S. wholesale price index.

Figure B.2: U.S. Imports and Exports as a Share of GDP



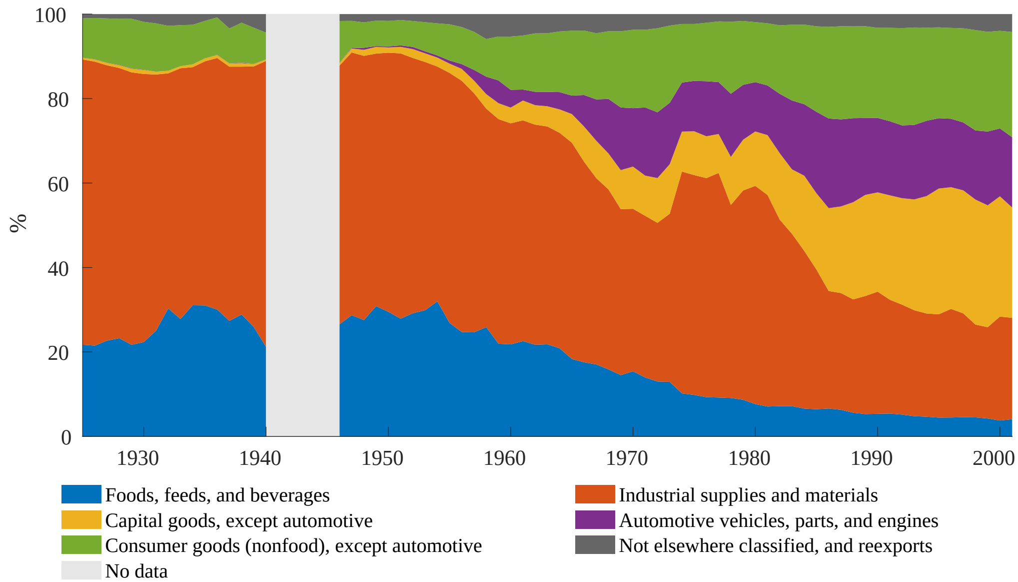
Notes: The blue line shows imports of goods as a share of GDP; the black dotted line shows exports of goods as a share of GDP.

for a non-negligible share. Exports similarly exhibit a large share of industrial supplies, alongside food products and capital goods. In the postwar period, the composition of trade shifts gradually. The importance of food and industrial supplies declines, while capital goods, automotive products, and consumer goods become increasingly prominent. These changes reflect the growing integration of manufacturing supply chains and the rising role of durable goods in international trade.

Figure B.5 zooms in on the distinction between crude and manufactured goods over a longer historical horizon. For imports, the composition is fairly stable over much of the sample, with crude materials, manufactured foods, semimanufactures, and finished manufactures each accounting for sizeable shares, and no clear monotonic trend until the later decades. For exports, by contrast, there is a gradual long-run shift away from crude goods toward manufactures.

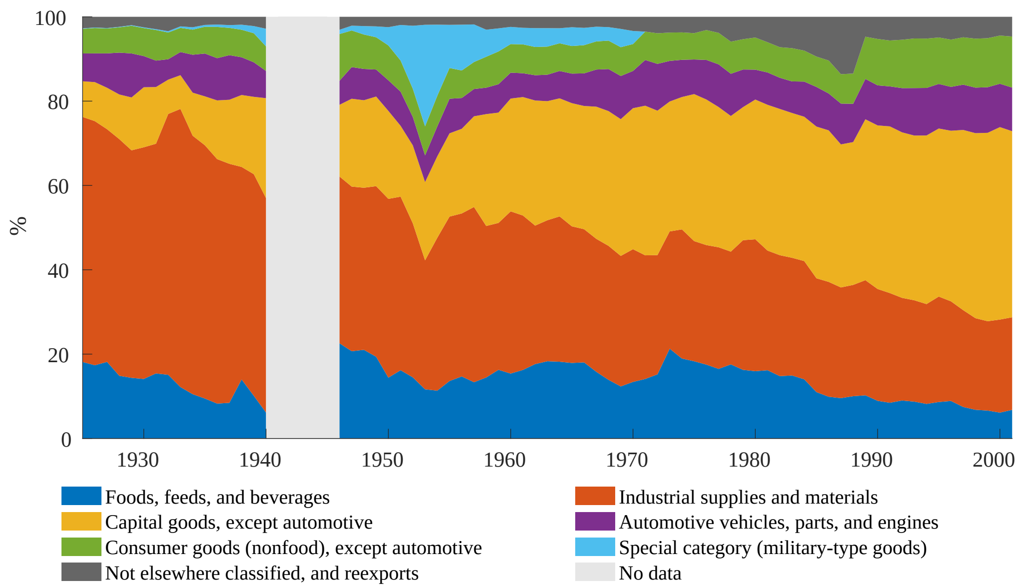
Overall, these descriptive patterns highlight two features relevant for our analysis. First, the United States has long been exposed to international trade, with exposure rising in the postwar era. Second, the composition of trade has evolved gradually, particularly after World War II, with a growing role for capital goods, consumer goods, and finished manufactures. These developments provide important context for interpreting the macroeconomic effects of tariff shocks over different historical periods.

Figure B.3: Composition of Imports by Broad End-Use Class



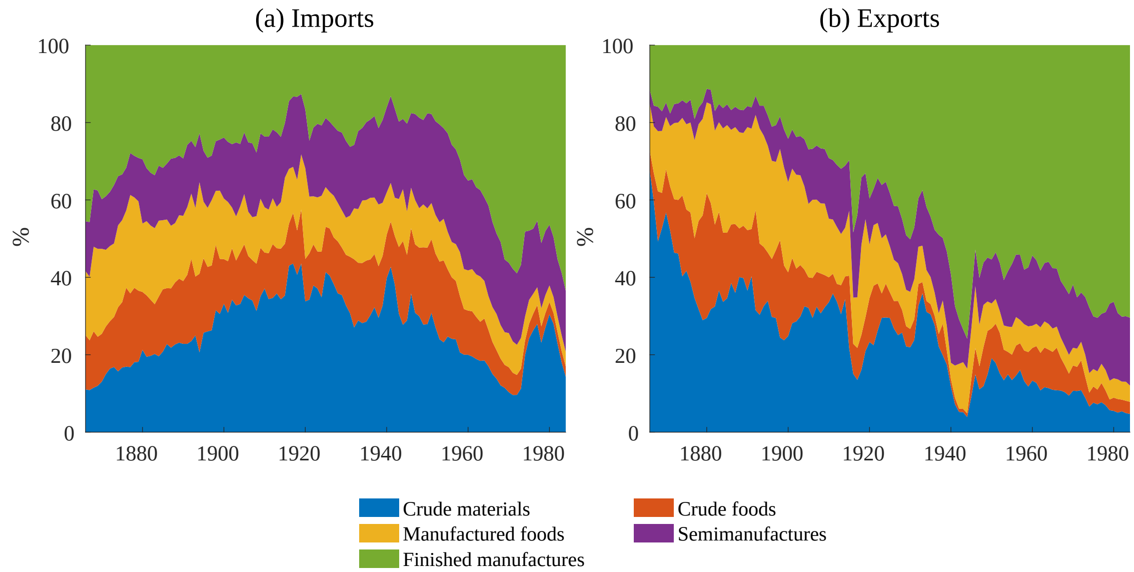
Notes: Areas represent each category's share of total U.S. imports by broad end-use class. Sample: 1925–2001. Gray shading indicates years with no data (1941-1945). Source: Irwin (2006), Table Ee506-520 in *Historical Statistics of the United States, Earliest Times to the Present*.

Figure B.4: Composition of Exports by Broad End-Use Class



Notes: Areas represent each category's share of total U.S. exports by broad end-use class. Sample: 1925–2001. Gray shading indicates years with no data (1941-1945). Source: Irwin (2006), Table Ee506-520 in *Historical Statistics of the United States, Earliest Times to the Present*.

Figure B.5: Composition of Imports and Exports: Crude and Manufactured Goods



Notes: Areas show each category's share of total U.S. imports in Panel (a) and exports in Panel (b). Sample: 1866–1984. Source: Irwin (2006), Table Ee446-457 in *Historical Statistics of the United States, Earliest Times to the Present*.

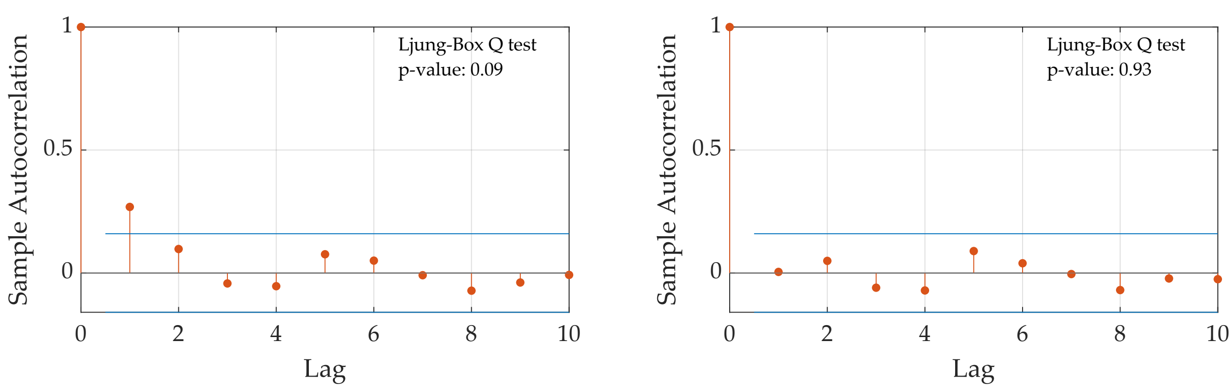
C. Instrument Diagnostics

We perform a series of validity checks on the instrument. Specifically, we examine the time-series properties of the exogenous tariff events by analyzing the autocorrelation and forecastability of the instrument. We also test a potential indirect channel by assessing whether economic conditions predict shifts to unified governments that enabled tariff legislation. Finally, we examine the instrument's correlation with other macroeconomic shocks from the literature.

Autocorrelation. Table 1 in the main text reports Granger causality tests assessing whether the narrative tariff shock is predictable from past macro-financial variables. The p-value on lagged values of the instrument is 0.15, indicating that, conditional on the included controls, its past realizations do not significantly predict current values. This provides supportive evidence that the instrument behaves like a serially uncorrelated shock. Figure C.1a presents the autocorrelation function of the instrument. Unconditionally, there is some evidence of autocorrelation: the first lag is significant, and the p-value of the Ljung–Box Q statistic – testing the joint null that autocorrelations up to 10 lags are zero – is 0.09.

Figure C.1: Autocorrelation Functions of the Tariff Shock Series

(a) Raw Tariff Shock Series (b) Purged Tariff Shock Series

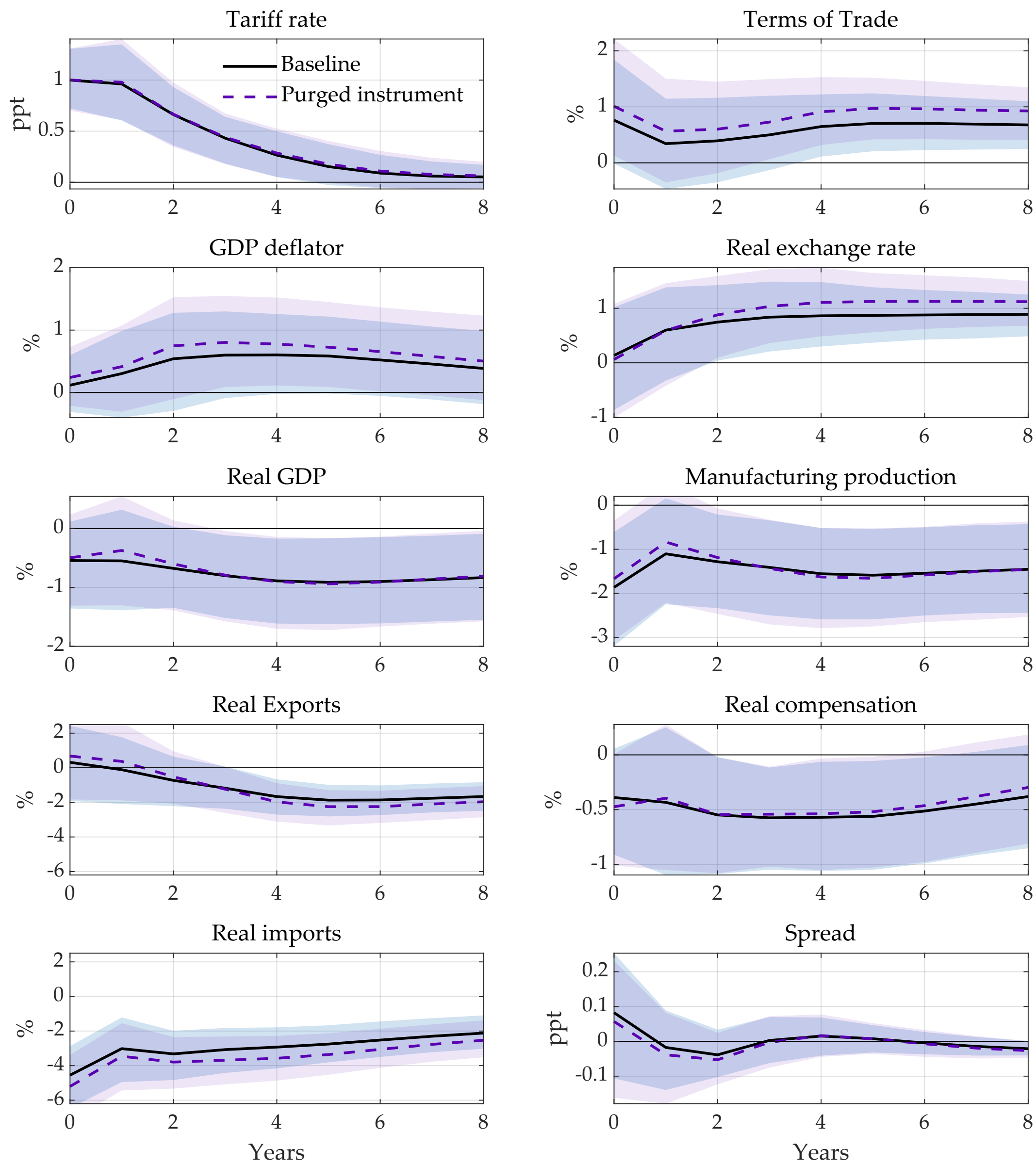


Thus, we follow the approach in Miranda-Agrippino and Ricco (2021) as a robustness check, and purge the distributed proxy, z_t, of predictable dynamics as an additional check. We regress the instrument on its lag and use the residual, $z_t^{\perp}$:

$$z_t = \phi_0 + \phi_1 z_{t-1} + z_t^{\perp}$$

Figure C.1b shows that the autocorrelation of the purged instrument $z_t^{\perp}$ is close to zero at all lags, confirming that any remaining serial correlation has been effectively removed. We also assess whether using the purged series, rather than the raw series, as the external instrument in the baseline VAR affects the estimated responses. As shown in Figure C.2, the point estimates and confidence bands are nearly identical, providing additional support for the validity of the instrument.

Figure C.2: Results for the Purged Instrument



Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated based on the external-instrument VAR model (3). Lag order: 2. Solid black line: baseline point estimate. Dashed purple line: point estimate for AR(1)-purged instrument. Light shaded areas: 90 percent confidence bands based on moving-block bootstrap.

Forecastability of political determinants. As documented in the paper, we do not find evidence that macroeconomic or financial variables have predictive power for our narrative tariff shock series. While this direct test is reassuring, we also examine a potential indirect channel: the political determinants of tariff policy. Historically, tariff legislation reflected a deep partisan divide. As a result, tariff changes were often driven by long-standing ideological commitments following a shift in power, rather than short-run economic conditions. Consistent with this view, all congressional tariff acts in our baseline sample occurred under a unified government, which is when the same party controls the House, Senate, and presidency. Irwin (1998) argues that unified government was effectively a necessary condition for the passage of major tariff legislation before 1934. This political context mitigates concerns that tariff reforms were systematically timed to the business cycle.

Table C.1: Forecastability of Unified Governments

Variable	p-value
Unified government	0.08
Tariffs and trade	
Tariff rate	0.63
Terms of trade	0.42
Real exports	0.69
Real imports	0.82
Macroeconomic variables	
Real GDP	0.66
Manufacturing production	0.69
Real compensation	0.96
GDP deflator	0.18
Producer price index	0.52
Financial variables	
Interest rate spread	0.88
REER	0.19
Real energy commodity price index	0.21
SP500	0.50
Government finances	
Real government debt	0.10
Real government revenue	0.40
Real government expenditure	0.25
Joint	0.69

Notes: p-values from Granger causality tests where the dependent variable is a dummy equal to 1 in the first year a government transitions from divided to unified control.

Nevertheless, economic conditions could still affect the *timing* of tariff reforms indirectly if they influence the likelihood of a unified government under which legislation is enacted. To assess this possibility, we repeat the Granger causality tests from Table 1, but instead test whether the variables have predictive power for transitions to unified government. Table C.1 shows that transitions to unified control exhibit some serial correlation. Importantly, the macroeconomic and financial variables do not systematically predict transitions to unified control, which further supports the exogeneity of the timing of tariff legislation in our sample.

Next, we examine whether recession episodes affect the occurrence of unified governments. Using NBER recession indicators, we estimate a probit model where the dependent variable is a dummy equal to 1 in the first year a government transitions from divided to unified control, and the independent variables are lagged recession dummies. Columns 1-2 of Table C.2 report results for the full baseline sample (1866–2024), while Columns 3-4 restrict the sample to 1866–1934, when Congress held exclusive authority over tariff legislation. Across all specifications, the estimates indicate that recessions do not significantly raise the probability of transitions to unified governments, consistent with our previous exercise.

Table C.2: Probit Model Results: Unified Government and Recessions

Unified Government	(1)	(2)	(3)	(4)
Recession dummy$_{t-1}$	0.137	0.085	-0.191	-0.149
	(0.271)	(0.278)	(0.417)	(0.428)
Recession dummy$_{t-2}$		0.264		0.635
		(0.278)		(0.516)
Constant	-1.341	-1.446	-1.068	-1.579
	(0.187)	(0.222)	(0.338)	(0.559)
Observations	159	159	69	69
Log Likelihood	-51.8	-51.3	-24.7	-23.8

Notes: Standard errors in parentheses. Columns 1–2 report results for the full baseline sample (1866–2024); Columns 3–4 restrict the sample to 1866–1934, when Congress held exclusive authority over tariff legislation.

Forecastability of tariff shocks: 1866–1934 subsample. As an additional robustness check, we repeat the forecastability tests of the narrative tariff shocks using only the 1866–1934 subsample, when tariff policy was set directly by Congress and constituted a highly salient political issue. The results in Table C.3 mirror those for the full sample: macroeconomic and financial variables do not predict the shocks, supporting their exogeneity in this period.

Table C.3: Forecastability of Exogenous and Endogenous Tariff Shocks (1866–1934)

	Narrative tariff shocks	
Variable/p-value	Exogenous	Endogenous
Tariffs and trade		
Instrument	0.07	0.03
Tariff rate	0.86	0.18
Terms of trade	0.95	0.90
Real exports	0.64	0.65
Real imports	0.42	0.35
Macroeconomic variables		
Real GDP	0.71	0.68
Manufacturing production	0.89	0.21
Real compensation	0.93	0.76
GDP deflator	0.80	0.65
Producer price index	0.73	0.95
Financial variables		
Interest rate spread	0.82	0.35
REER	0.66	0.88
Real energy commodity price index	0.77	0.85
SP500	0.75	0.33
Government finances		
Real government debt	0.97	0.48
Real government revenue	0.55	0.29
Real government expenditure	0.94	0.19
Joint	1.00	0.00

Notes: p-values of a series of Granger causality tests of the narrative tariff shock series using a selection of macroeconomic and financial variables.

Correlation with other shocks. Next, we examine whether our instrument is correlated with other shocks used in the literature. We briefly summarize the construction of the shocks reported in Table C.4. The Economic Policy Uncertainty (EPU) shock is obtained from a VAR by ordering the EPU index from Baker, Bloom, and Davis (2016) first and using a Cholesky decomposition. We construct a shock to real energy commodity prices in the same way, ordering the energy price series first in the VAR.

The fiscal policy shocks are taken from Romer and Romer (2010) and Ramey and Zubairy (2018). Romer and Romer (2010) identify exogenous tax changes motivated by long-run policy objectives rather than current economic conditions. Ramey and Zubairy (2018) construct military spending news shocks based on narrative evidence of major defense policy announcements that signal unanticipated changes in expected future government spending. As both series are originally quarterly, we assign each observation to the first month of the corresponding quarter and aggregate them to the annual frequency using the same procedure as for our distributed proxy.

Finally, we consider monetary policy shocks based on the monthly narrative dates identified by Friedman and Schwartz (1963) and Romer and Romer (2023), which capture episodes in which policy actions were taken independently of contemporaneous or expected real-economic conditions and are therefore treated as plausibly exogenous. We assign each event to its respective month, construct a series of signed dummy variables, and then aggregate these to the annual frequency using the same procedure applied to our distributed proxy.

Table C.4 shows that none of the structural shocks considered are significantly correlated with our instrument. All correlation coefficients are below 0.15, and the associated p-values are well above conventional levels of statistical significance.

Table C.4: Correlation with Other Shock Measures

Shock	Source	ρ	p-value	n	Sample
Economic policy uncertainty	Baker, Bloom, and Davis (2016)	0.11	0.24	123	1900–2024
Tax policy	Romer and Romer (2010)	-0.06	0.66	63	1945–2007
Military spending news	Ramey and Zubairy (2018)	-0.13	0.15	126	1890–2015
Monetary policy	Friedman and Schwartz (1963); Romer and Romer (2023)	0.13	0.22	97	1920–2016
Real energy commodity prices	Jacks (2019); Wright (2006)	-0.06	0.47	157	1866–2024

Notes: Correlation coefficients of the narrative instrument series with a wide range of different shock measures from the literature, including fiscal, monetary and oil shocks. ρ is the Pearson correlation coefficient, the p-value corresponds to the test whether the correlation is different from zero and n is the sample size.

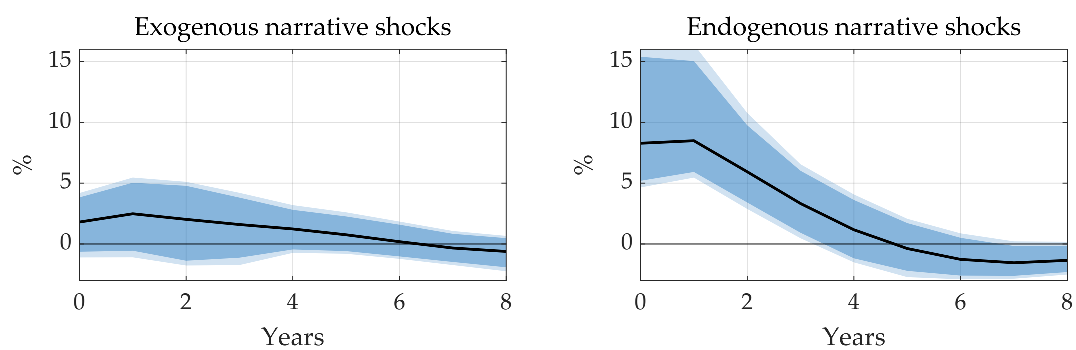
D. Additional Results

In this appendix, we report some additional results, complementing the analysis in the paper.

D.1. Endogenous vs. Exogenous Shocks

In this appendix, we further compare the dynamic effects of tariff changes classified as endogenous and exogenous based on our narrative approach. We focus in particular on the response of government spending, as this provides a direct test of the underlying motivation for tariff changes. Figure D.1 reports the corresponding impulse responses. Endogenous tariff episodes – enacted in response to prevailing macroeconomic or fiscal conditions – are followed by a significant increase in government spending, consistent with their revenue-raising or fiscal motivation. By contrast, no comparable response is observed following plausibly exogenous tariff shocks. These findings highlight the importance of accounting for endogeneity in tariff changes when estimating their effects.

Figure D.1: The Effect on Government Spending



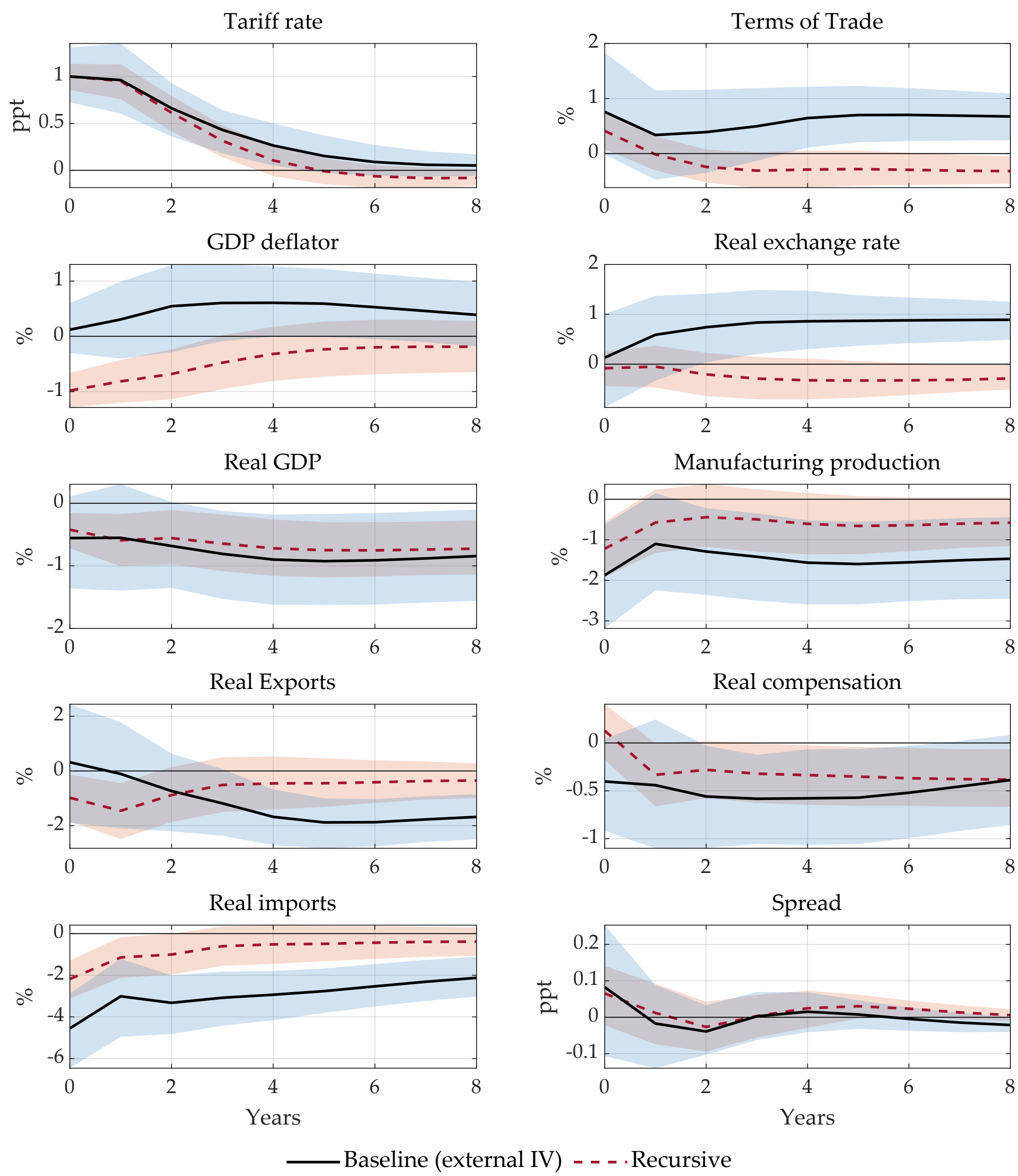
Notes: Government spending response to endogenous and exogenous tariff shocks, estimated based on the external-instrument VAR model (3). Both shocks are normalized to increase the dutiable tariff rate by 1 percentage point on impact. Solid line: point estimate. Dark and light shaded areas: 90 and 95 percent confidence bands based on moving-block bootstrap.

D.2. Comparison with Recursive Approach

A commonly used identifying assumption in the literature is that tariff changes are predetermined with respect to the macroeconomy (e.g., Furceri et al., 2022; Franconi and Hack, 2025). Under this approach, innovations to the tariff rate are treated as exogenous shocks. However, this assumption is unlikely to hold in practice. It mixes endogenous and exogenous tariff changes and does not account for non-policy-induced variation in observed tariff rates arising from behavioral responses.

To assess the importance of these issues, we compare our external-instrument approach to a recursive VAR in which the average tariff rate is ordered first. Figure D.2 presents the resulting impulse responses.

Figure D.2: SVAR-IV vs. Recursive Identification



Notes: Impulse responses to a tariff shock identified based on the baseline external-instrument VAR compared to a tariff shock identified using a recursive VAR. Lag order: 2. Solid line: point estimate. Dark and light shaded areas: 90 and 95 percent confidence bands based on moving-block bootstrap.

The responses differ markedly from our baseline estimates. In the recursive specification, the responses of imports and exports are muted with exports falling immediately on impact, and the exchange rate depreciates rather than appreciating. The most pronounced difference arises in the price response: the GDP deflator declines significantly and persistently. These differences reflect the fact that the recursive approach fails to isolate exogenous tariff shocks.

Overall, these results underscore the importance of accounting for the endogeneity of tariff changes. Our narrative classification and external-instrument strategy isolate plausibly exogenous tariff shocks and yield qualitatively different and economically meaningful estimates of their effects on trade, output, and prices.

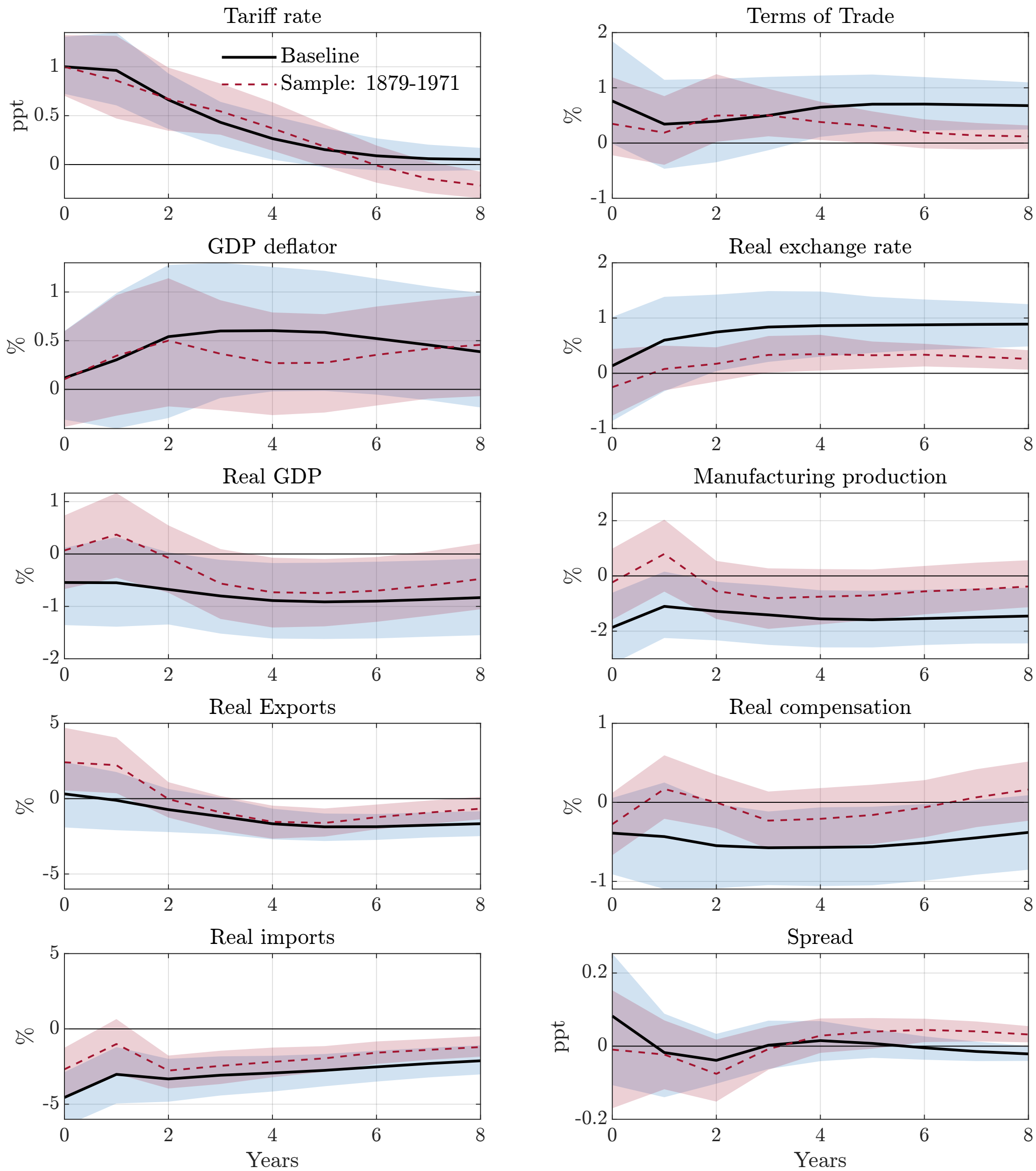
D.3. Comparison to Barnichon and Singh (2025)

Our empirical results regarding the effects of tariffs on real activity are broadly consistent with Barnichon and Singh (2025), with minor differences driven by a different empirical setup. Barnichon and Singh (2025) (i) consider a specification in first-differences instead of levels, (ii) use a smaller set of exogenous tariff dates, (iii) use a smaller set of controls. Regarding the response of inflation, the results are consistent in the post-World War II sample – a tariff increase leading to a decline in inflation – but the results do differ quantitatively in the pre-World War II sample. Unlike in the present paper where the inflation response is imprecisely estimated and not statistically different from zero, Barnichon and Singh (2025) document a sharp drop in inflation after a tariff increase. The chief reason for this difference is the existence of residual endogeneity in the Barnichon and Singh (2025) tariff shocks coming from tariff policy responding to price developments. Most tellingly, the Emergency tariff of 1921 was classified as an exogenous change by Barnichon and Singh (2025), although it was meant as a temporary protection to farmers from *falling prices*. The residual endogeneity of this large tariff shock leads to a marked decline in inflation. Indeed, removing this date from the Barnichon and Singh (2025) shock series removes most of the inflation drop. Using Barnichon and Singh (2025)'s exact specification, the inflation responses estimated over pre-World War II data are similarly uncertain and not different from zero – once this episode is excluded – whether using our present set of exogenous shocks or the shock series used by Barnichon and Singh (2025).

D.4. Additional Results on Monetary System and Foreign Responses

Responses under gold-backed monetary system. To shed light on the role of monetary policy, we estimate our model on the sample featuring gold-backed monetary systems (1879–1971). Figure D.3 compares the responses over this sample to our baseline responses.

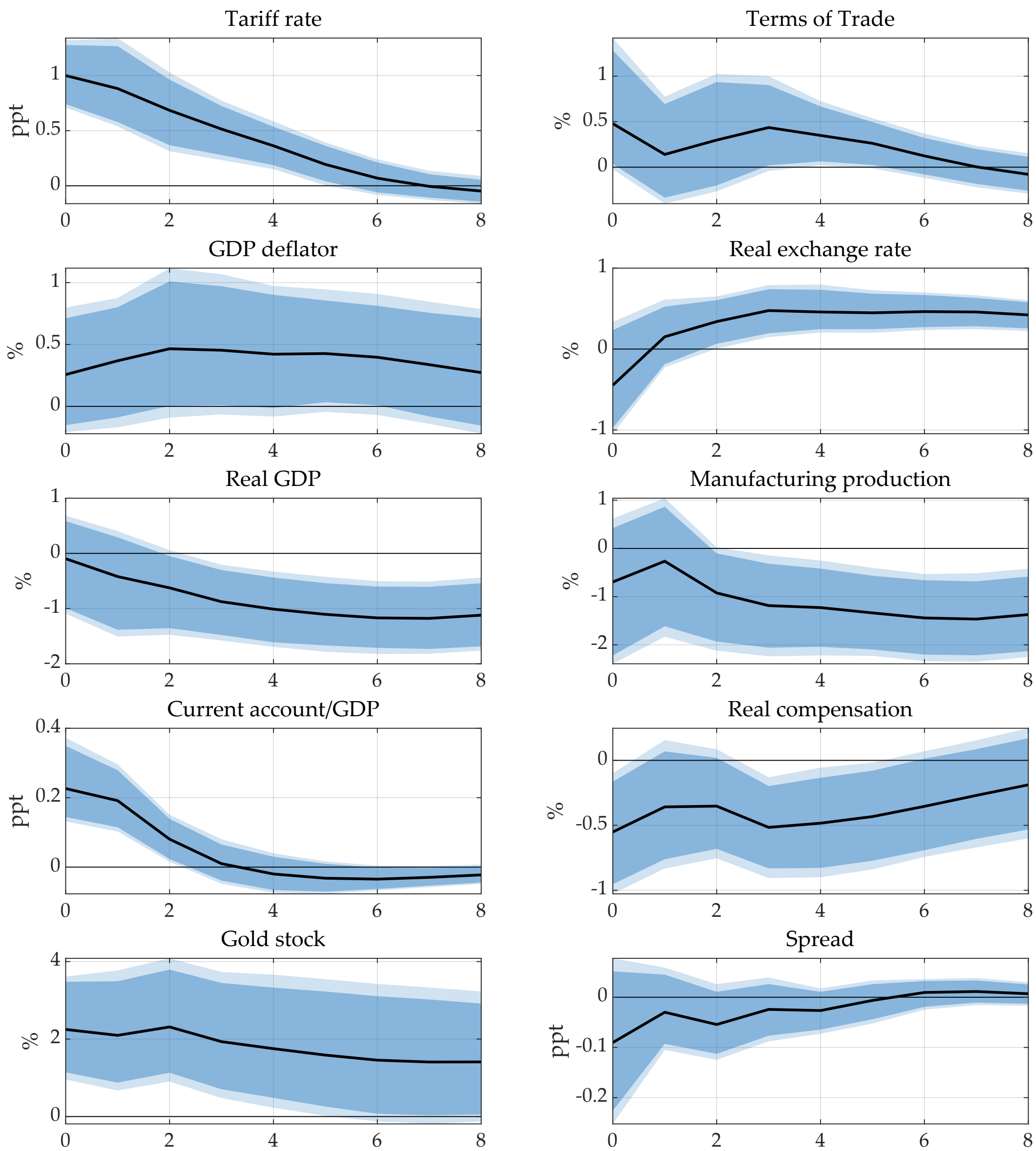
Figure D.3: Responses for Era of Gold-Backed Monetary Systems vs. Full Sample



Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated based on the external-instrument VAR model (3) over different sample periods. Lag order: 2. Solid black line: baseline responses. Dashed red line: responses for sample 1879–1971. Light shaded areas: 90 percent confidence bands based on moving-block bootstrap.

Because we are interested in the response of the stock of gold, we also estimate an alternative specification that includes the current account and the gold stock in lieu of imports and exports. In the main text, we show a selection of these responses. For completeness, we reproduce the full set of responses below in Figure D.4.

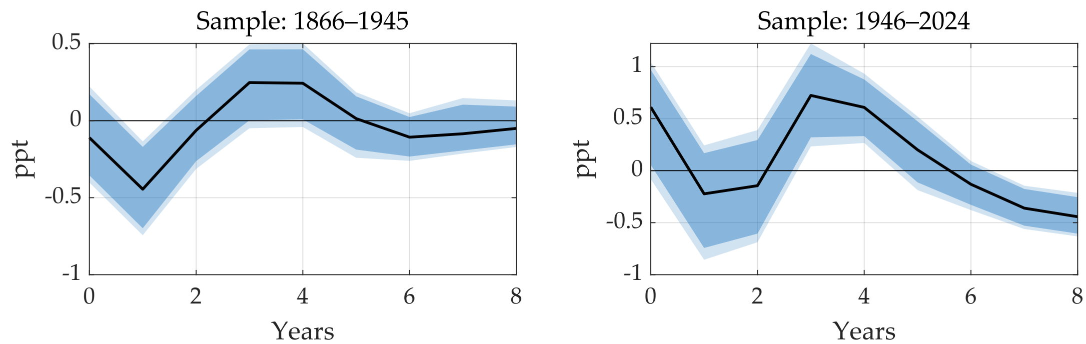
Figure D.4: Responses for Era of Gold-Backed Monetary Systems, Alternative Spec.



Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated based on the external-instrument VAR model (3), including the current account over GDP and the stock of gold in lieu of the import and export variables. Sample: 1879–1971. Lag order: 2. Solid black line: point estimate. Dark and light shaded areas: 90 and 95 percent confidence bands based on moving-block bootstrap.

Monetary stance pre- and post-World War II. We use the real short-term interest rate as a proxy for the monetary policy stance, recognizing that this measure is imperfect, particularly in the earlier period. In the pre-World War II sample, the real short rate declines following a tariff shock, consistent with a more accommodative monetary response under gold-backed monetary arrangements. By contrast, in the postwar period, the real short rate rises temporarily, indicating a more contractionary policy response. This shift in the behavior of the real rate points to a markedly different monetary adjustment across regimes.

Figure D.5: Real Short-Term Rate Responses



Notes: Impulse responses of real short-term interest rates to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated using the marginal VAR approach. Solid line: point estimate. Dark and light shaded areas: 90 and 95 percent confidence bands based on moving-block bootstrap.

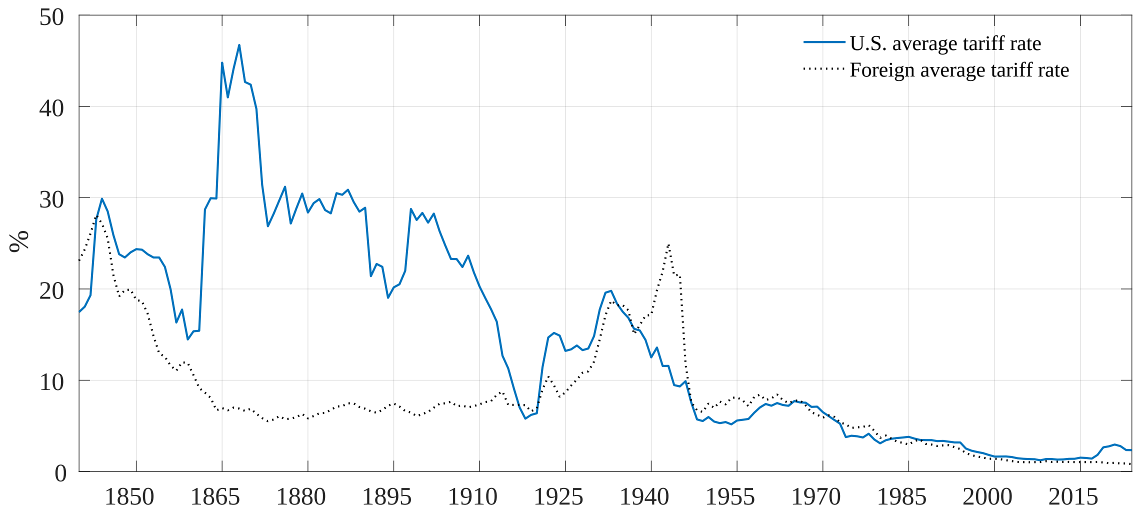
More on foreign policy responses. Figure D.6 shows the U.S. average tariff rate relative to total imports (blue line) and the export-weighted foreign tariff rate (black dotted line) from 1840 to 2024. The foreign tariff rate is constructed as the average tariff rate relative to total imports across thirteen major trading partners.

As shown in Figure D.6, the U.S. average tariff rate and the export-weighted foreign tariff rate were at similar levels in 1840. U.S. tariffs increased rapidly during the Civil War and remained elevated for several decades after. In contrast, foreign tariff rates gradually declined over this period and remained below 10 percent on average, showing no evidence of retaliation in response to U.S. tariff hikes.

This pattern changed during the interwar period, when both U.S. and foreign tariff rates rose in the early 1920s and 1930s. The 1930 Smoot-Hawley Tariff Act in particular resulted in widespread international retaliation. Following the passage of the Reciprocal Trade Agreements Act (RTAA), U.S. tariffs gradually declined, while foreign tariff rates continued to increase until the end of the Second World War. In the postwar period, both U.S. and foreign tariffs declined gradually, reflecting multilateral tariff reductions under the eight GATT rounds.

Overall, Figure D.6 indicates that U.S. tariff policy was largely unilateral prior to the 1920s, with little evidence of reciprocity, whereas in the modern period U.S. and foreign

Figure D.6: Average U.S. and Export-Weighted Foreign Tariff Rates, 1840–2024



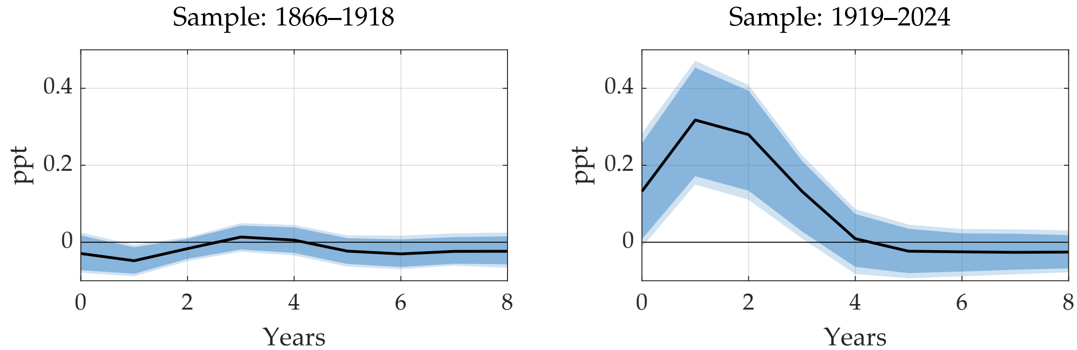
Notes: The chart shows the average U.S. tariff rate and the export-weighted average foreign tariff rate from 1840 to 2024 (percent). The blue line represents the U.S. tariff rate, measured as duties relative to total imports. The black dotted line represents the foreign tariff rate, calculated as an export-weighted average of tariffs imposed by major trading partners. These countries include the United Kingdom, France, Germany, Italy, Canada, Switzerland, Belgium, the Netherlands, Portugal, Spain, Japan, Mexico, and Brazil.

tariff rates exhibit greater comovement. Figure 9 in the main text corroborates this pattern: the export-weighted foreign tariff rate does not respond significantly in the pre-World War II period, but rises significantly and persistently following a U.S. tariff shock in the post-World War II period.

We repeat the exercise for the subsamples 1866–1918 and 1919–2024 and report the results in Figure D.7. The difference is now even somewhat starker than in Figure 9. In the earlier period, the response is close to zero and insignificant. In contrast, in the later period, foreign tariffs increase significantly and persistently following a U.S. tariff shock. The weaker response in 1866–1918 likely reflects the episodes of retaliation in the early 1920s and 1930s.

Taken together, the evidence presented in this appendix suggests that systematic retaliation was largely absent in the earlier part of the sample, while retaliation and coordination became more prevalent in the modern period. This pattern is consistent with more pronounced contractionary effects of tariffs in the postwar period and with a transmission mechanism that operates more strongly through exports and aggregate demand.

Figure D.7: Responses of Foreign Tariff Rates



Notes: Impulse responses of foreign tariff rates to a U.S. tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact. The foreign tariff rate is defined as an export-weighted average of trading partners' tariffs. Left panel: 1866–1918. Right panel: 1919–2024. Responses are estimated using the VAR model (3), augmented with the foreign tariff rate. Lag order: 2. Solid line: point estimate. Dark and light shaded areas: 90 and 95 percent confidence bands based on moving-block bootstrap.

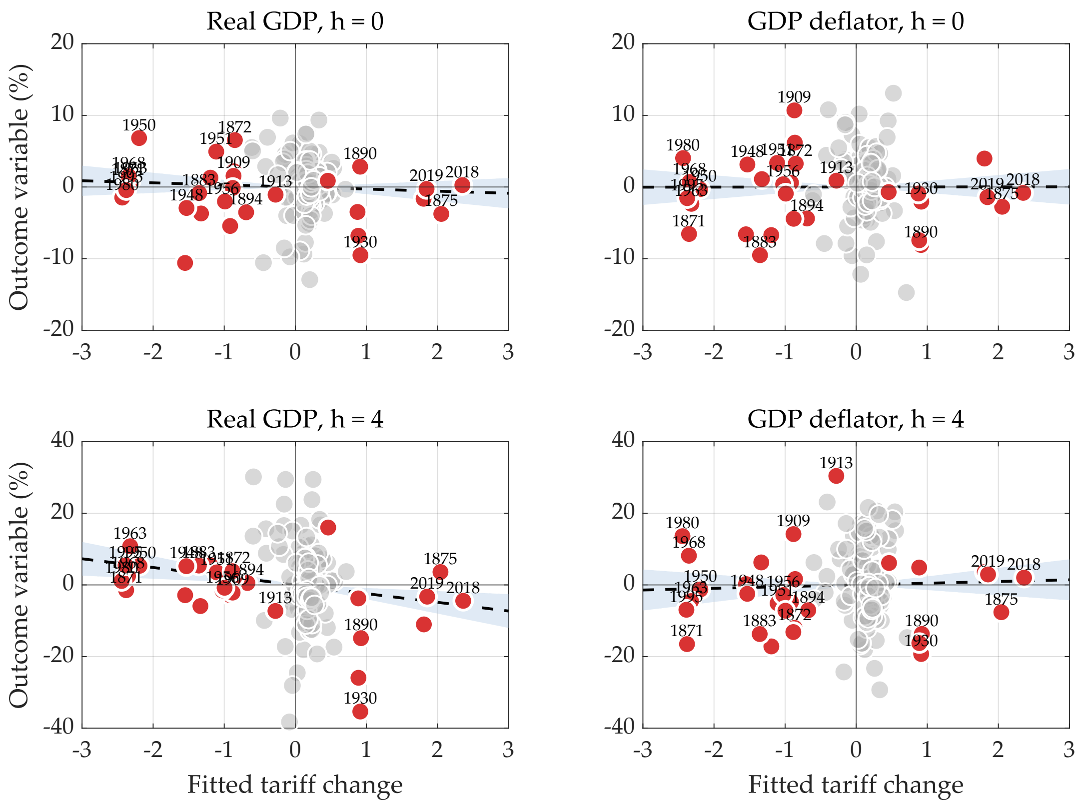
D.5. Additional LP-IV Results

In this appendix, we visualize the identifying variation underlying the LP-IV specification, akin to the event-study analysis in Section 3.3. In addition, we use the LP-IV specification to test for pre-trends.

LP-IV scatterplot. Figure D.8 shows scatter plots of real GDP and the deflator, on impact and four years ahead, against the instrumented dutiable tariff rate, residualized with respect to the controls. The plots correspond to the LP-IV estimates in Figure 10a. There is no evidence that the relationship between tariffs and real GDP is driven by a small number of outliers.

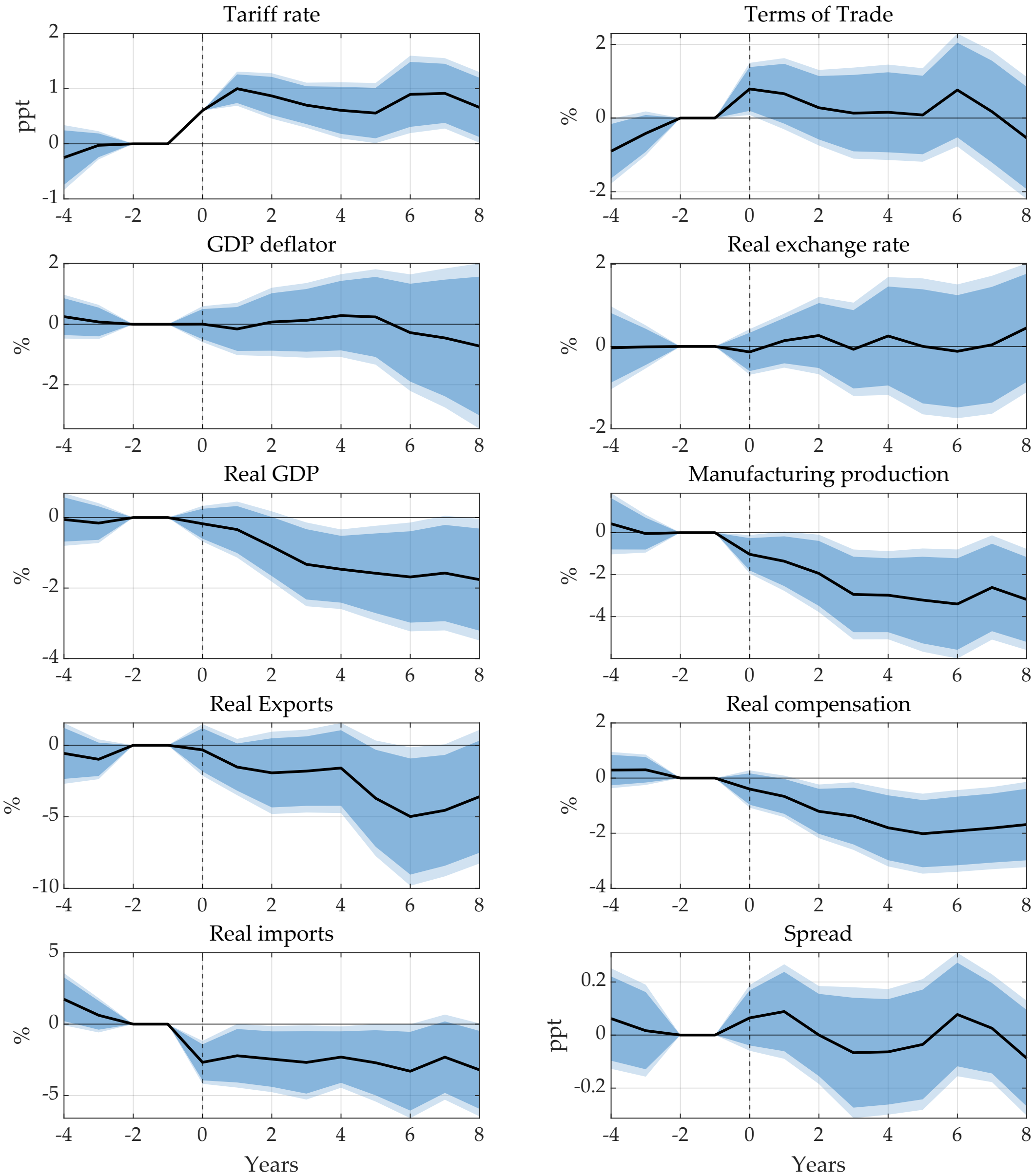
Pre-trends. The Granger causality test in Table 1 already shows that the narrative tariff shock series is not systematically anticipated by observable economic conditions. To further support this identification assumption, we perform a pre-trend test within the LP-IV framework. Figure D.9 shows that there are no economically meaningful or statistically significant pre-trends. Note that the pre-trends from $h = -(p+1), \ldots, -1$ are zero by construction because of the lags included in our local projections.

Figure D.8: Narrative Tariff Shocks and the Economy Full Sample



Notes: Scatter plot illustrating the LP-IV relationship between the residualized fitted shock and the cumulative change in outcome variables at different horizons h. As outcome variables, we consider the change in real GDP and the GDP deflator, in the current year and over a four year horizon. Both variables are residualized with respect to our baseline controls. Sample: 1866–2024. Lag order: 1. Shaded areas represent 90 percent confidence bands based on heteroskedasticity-robust standard errors.

Figure D.9: Testing for Pre-trends



Notes: Impulse responses to a tariff shock, estimated based on the LP-IV approach. Solid line: point estimates. Shaded areas represent 90 and 95 percent confidence bands based on heteroskedasticity-robust standard errors. The responses are normalized to match the peak effect in the baseline external-instrument VAR.

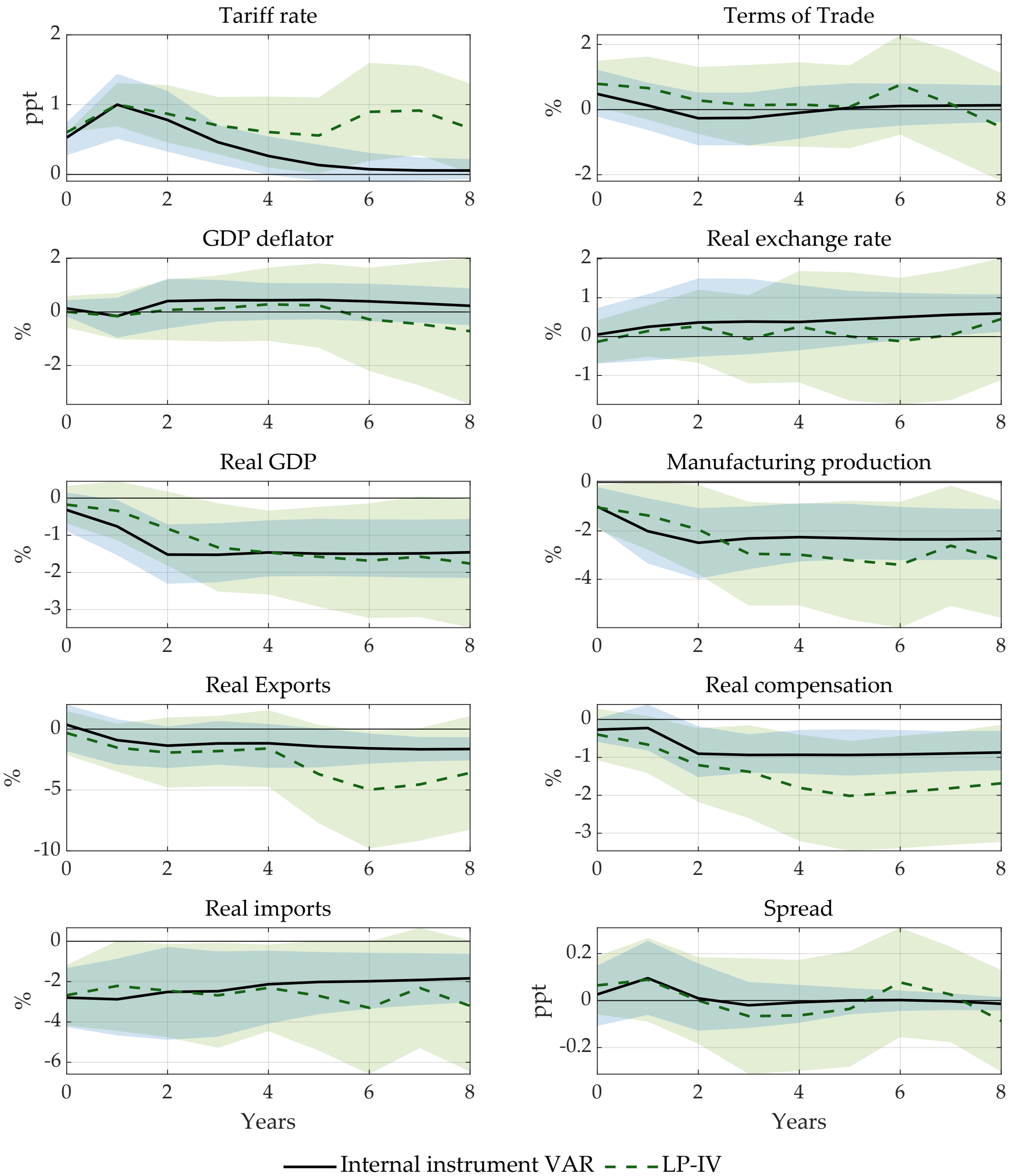
E. Sensitivity Analysis

This appendix presents an extensive robustness analysis to ensure that the empirical results are not driven by specific event selections, identifying assumptions, model specification choices, or data sources.

E.1. Relaxing VAR Assumptions

In this section we present the full set of responses for the specification in Figure 10a. As discussed in the main text, our external-instrument VAR relies on assumptions such as invertibility and limited lag-truncation bias. To test robustness, we re-estimate tariff shock effects using alternative methods. First, we apply the internal-instrument approach of Plagborg-Møller and Wolf (2019), which relaxes invertibility while preserving VAR dynamics. Second, we estimate an LP-IV specification that further relaxes dynamic restrictions. Figure E.1 shows that both approaches yield impulse responses very similar to the baseline VAR. The main difference is lower precision, with wider confidence bands.

Figure E.1: Relaxing VAR Assumptions



Notes: Impulse responses to a tariff shock estimated based on alternative models: the internal instrument VAR and the LP-IV approach. Lag order: 1 in LP-IV and 2 in VAR. Solid and dashed lines: point estimates. Shaded areas: 90 percent confidence bands based on moving-block bootstrap (internal instrument VAR) and heteroskedasticity-robust standard errors (LP-IV). Responses are normalized to match the peak effect in the baseline external instruments VAR.

E.2. Event Selection and Instrument Construction

A core element of our identification strategy is the construction of a narrative instrument based on exogenous tariff events. In Appendix C, we already showed that the resulting shock series exhibits many of the desired properties for a valid instrument. In this appendix, we present a series of robustness checks concerning the selection of events as well as choices regarding the timing and size of the shock.

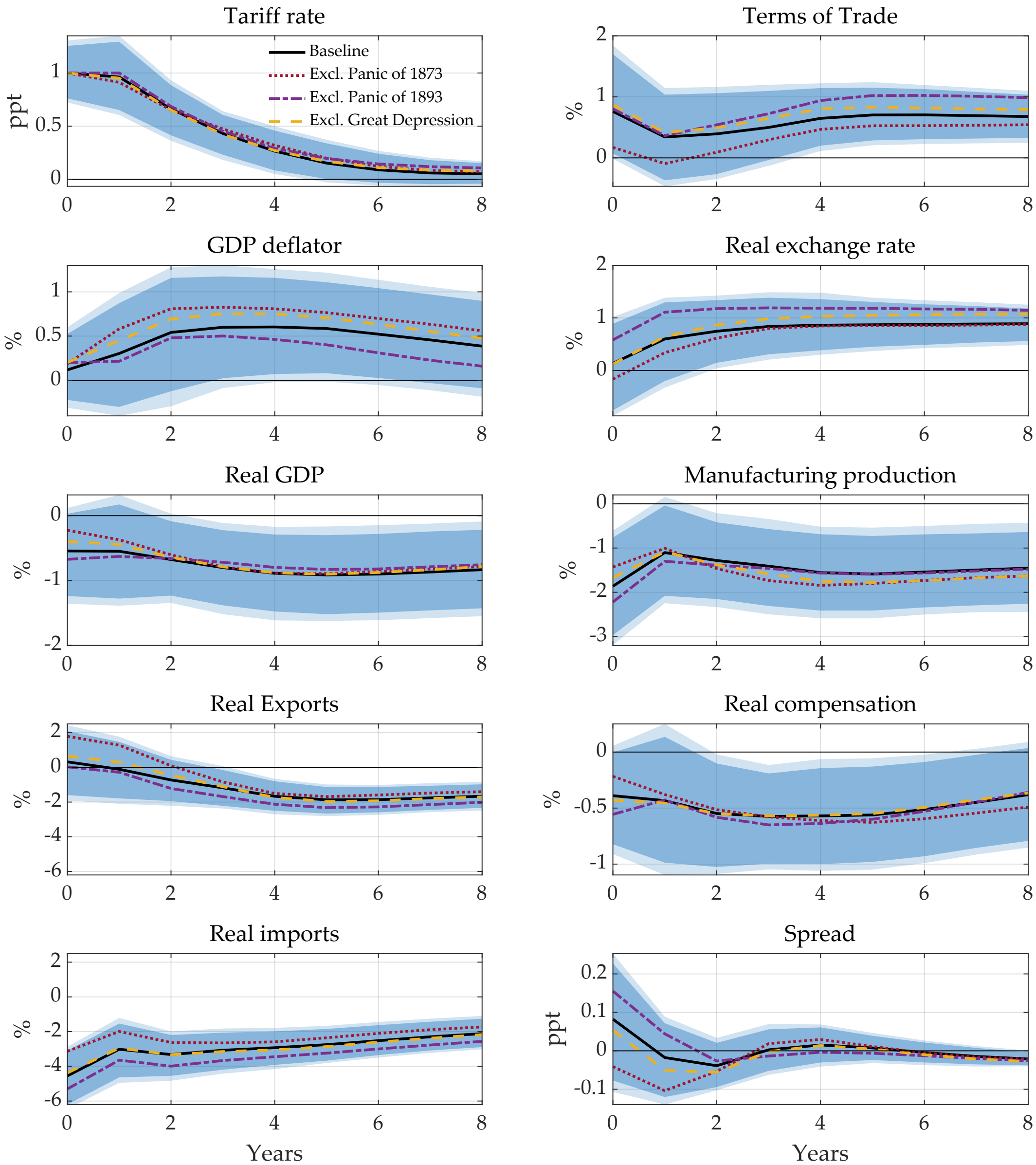
E.2.1. Accounting for Potential Confounders

A potential concern is the presence of confounding shocks. Even when tariff changes are plausibly exogenous in terms of their motivation, they may, in finite samples, coincide with other macroeconomic disturbances – such as financial crises or recessions – that independently affect the outcome variables of interest. Such coincident shocks could bias the estimated responses and complicate the interpretation of the results as reflecting the causal effects of tariff policy.

Thus, we examine whether our estimated responses are robust to excluding tariff episodes enacted in close proximity to major financial crises or recessions. Specifically, the Tariffs of 1870, 1872, and 1875 were implemented around the Panic of 1873, the McKinley Tariff of 1890 and the Wilson–Gorman Tariff of 1894 were enacted around the Panic of 1893, and the Smoot-Hawley Tariff of 1930 was implemented during the Great Depression. Figure E.2 shows that the estimated responses remain closely aligned with the baseline results when we exclude these events from our instrument and re-estimate our baseline model.

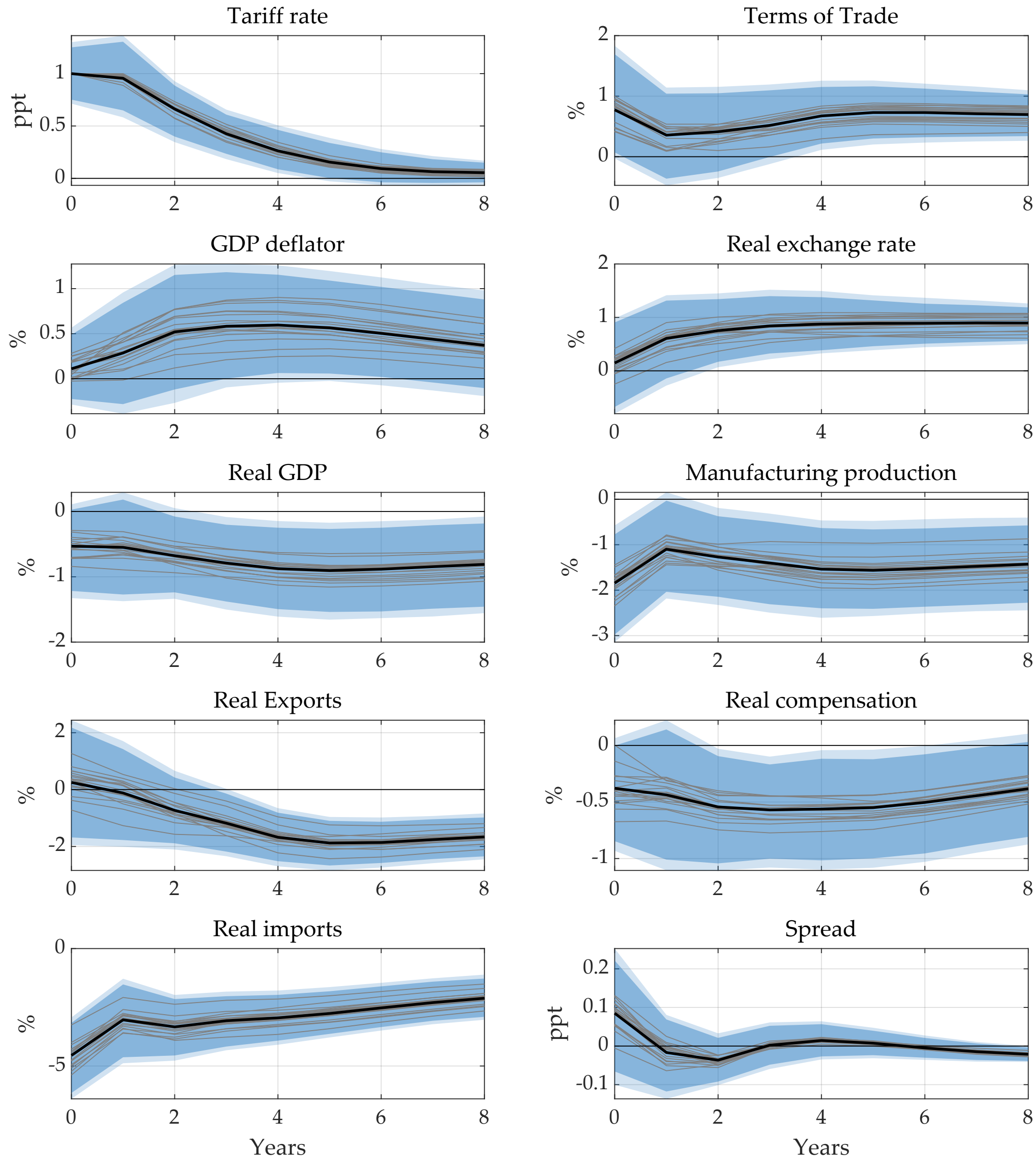
To assess the influence of specific events more systematically, we conduct a jackknife exercise. In each iteration, we remove one event from the instrument series and re-estimate the external-instrument VAR. Figure E.3 plots the baseline responses in black and the jackknife responses in gray. The results indicate that no single large tariff event drives the estimated responses. Depending on which event is excluded, the responses become slightly larger or somewhat attenuated. In most cases, the responses remain well within the 90 percent confidence bands of the baseline model.

Figure E.2: Excluding Events Around Major Recessions



Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated based on the external-instrument VAR model (3), excluding selected events from the instrument. Solid line and shaded areas: baseline response with 90 and 95 percent confidence bands based on moving-block bootstrap. Dashed and dotted lines: point estimates when excluding selected events from the instrument.

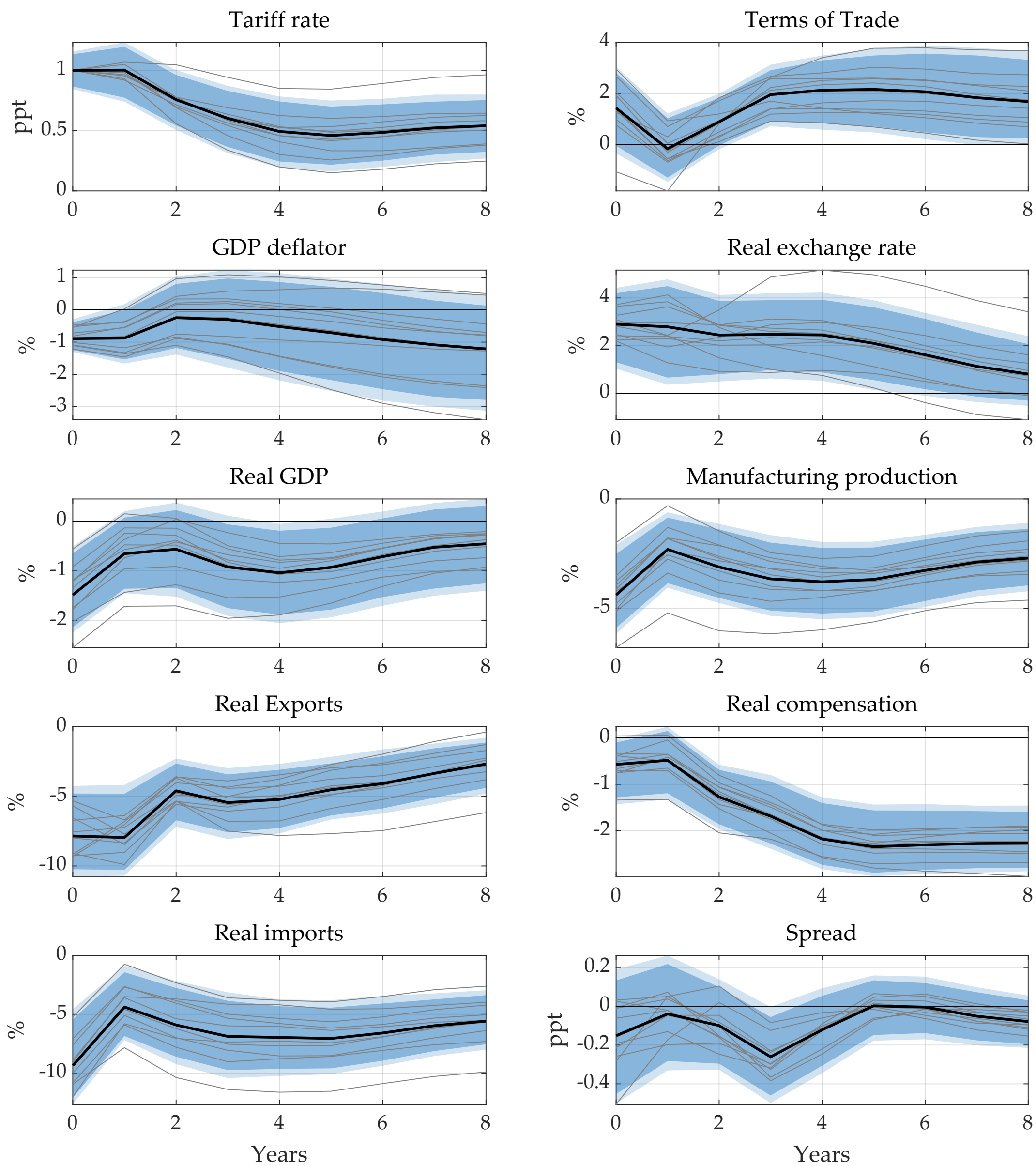
Figure E.3: Results from Jackknife Exercise



Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated based on the external-instrument VAR model (3), excluding one event in the instrument at a time. Sample: 1866–2024. Lag order: 2. Solid line and shaded areas: baseline response with 90 and 95 percent confidence bands based on moving-block bootstrap. Gray lines: jackknife responses, obtained by re-estimating the VAR after removing one tariff event from the instrument series at a time.

We repeat the jackknife exercise for the postwar sample and plot the results in Figure E.4. In contrast to the full historical sample, the responses are noticeably less stable across jackknife iterations, reflecting the limited number of major tariff events and the resulting sensitivity to the exclusion of individual episodes. This comparison underscores the key appeal of the longer historical sample used in the paper: the larger number of identified events reduces the influence of single episodes and leads to more stable estimated responses.

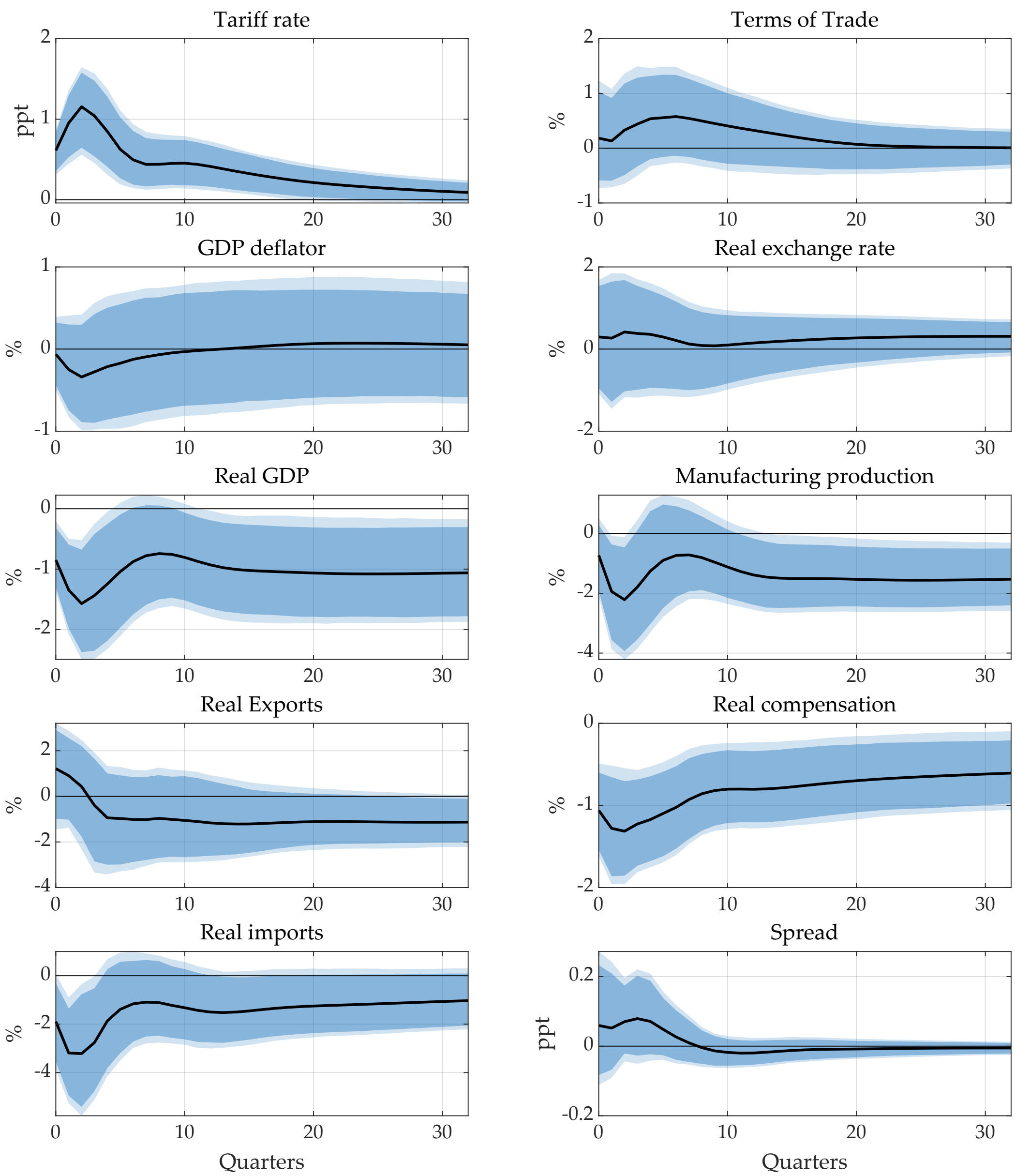
Figure E.4: Results from Jackknife Exercise for the Postwar Sample



Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated based on the external-instrument VAR model (3), excluding one event in the instrument at a time. Sample: 1946–2024. Lag order: 2. Solid line and shaded areas: baseline response with 90 and 95 percent confidence bands based on moving-block bootstrap. Gray lines: jackknife responses, obtained by re-estimating the VAR after removing one tariff event from the instrument series at a time.

Finally, we estimate a quarterly specification for the sample 1866Q1–2024Q4 to improve the timing of tariff shocks and mitigate concerns that annual tariff changes may be confounded with other contemporaneous developments. We present the main variables of interest in Figure 10c and present the full set of responses in Figure E.5. While the series are necessarily noisier, the resulting impulse responses closely mirror the baseline estimates.

Figure E.5: Responses From Quarterly VAR



Notes: Impulse responses to a tariff shock estimated based on the external-instrument VAR model (3). Sample: 1866Q1-2024Q4. Lag order: 2. Solid and dashed lines: point estimate. Dark and light shaded areas: 90 and 95 percent confidence bands based on moving-block bootstrap.

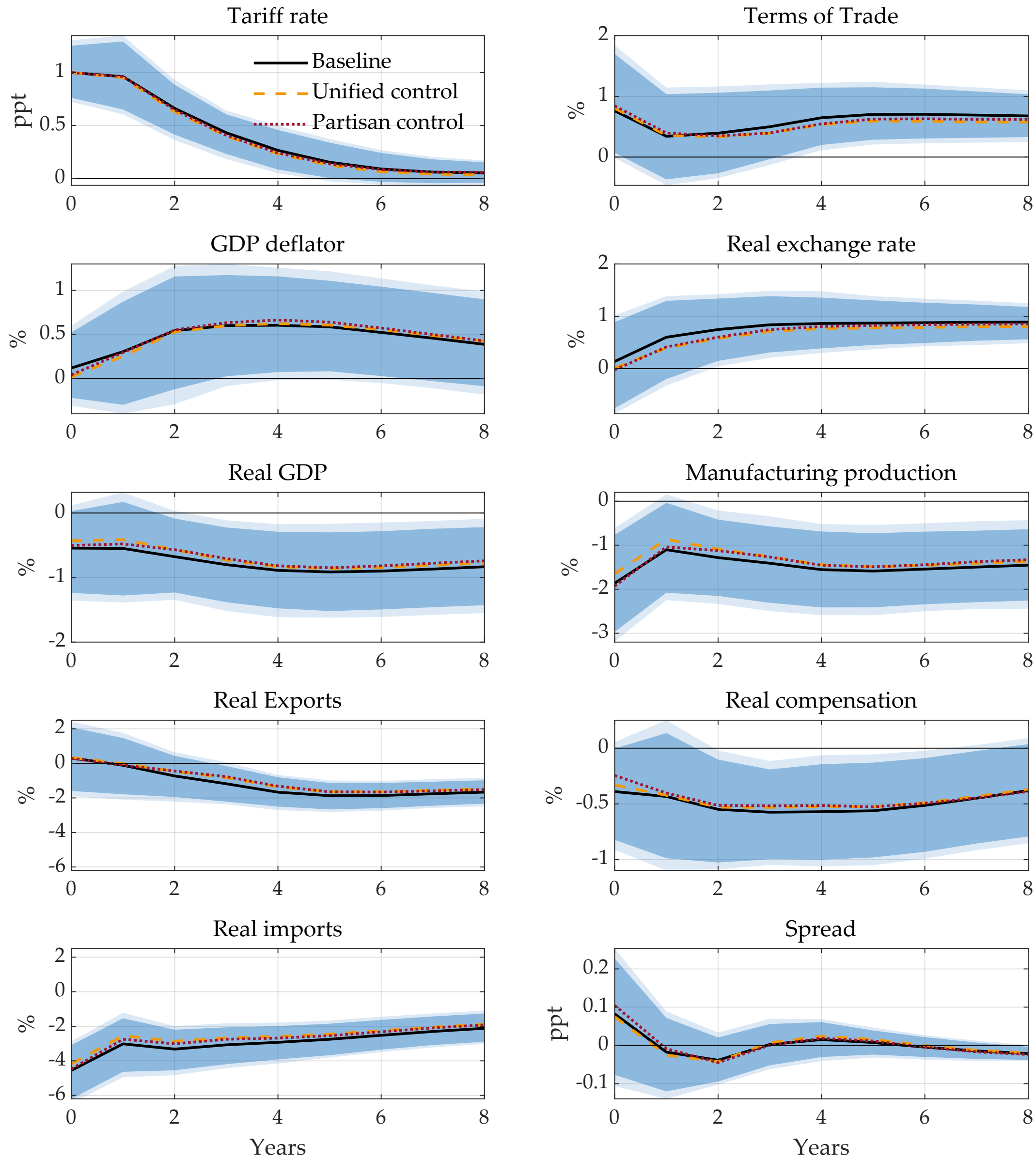
A second concern is that tariff changes may coincide with broader partisan policy shifts. As discussed in Appendix C, all congressional tariff acts in our baseline sample occurred under a unified government, which – due to the strong partisan divide on tariff policy – was often a necessary condition for the passage of major tariff legislation prior to 1934. This raises the possibility that these episodes coincide with the implementation of a broader set of partisan policies, so that the estimated responses may partly reflect the effects of wider policy packages associated with Republican or Democratic administrations.

To address this concern, we conduct two sets of robustness checks. First, in the narrative companion paper (Den Besten and Känzig, 2026), we exclude tariff changes that were implemented as part of larger legislative packages or accompanied by significant fiscal measures, and find that the results remain robust. Second, we augment the baseline specification with controls for unified government and partisan control. Figure E.6 reports the results. The yellow dashed line shows estimates including a single indicator for unified government, capturing periods of greater legislative capacity. The red dotted line instead distinguishes between unified Democratic and Republican control, allowing for differences across partisan administrations. Across both exercises, the estimated responses remain largely unchanged, suggesting that our results are not driven by coincident partisan policy shifts.

A related concern pertains to the narrative classification itself. By construction, the classification focuses on the primary motivation underlying each tariff change, but policy decisions often reflect multiple considerations. Secondary or contextual motivations – such as fiscal pressures or broader economic conditions – may have influenced the timing or design of tariff reforms, potentially introducing residual endogeneity. More generally, narrative classification necessarily involves judgment in weighing competing historical accounts.

We address this concern by assessing the sensitivity of our results to alternative classifications of borderline cases. In particular, we consider episodes in which fiscal considerations or other secondary motivations may have played a role and examine whether reclassifying these events affects the results. We also analyze cases in which economic conditions deteriorated during the legislative process or negotiations over tariff reform, potentially influencing the timing or content of the final law. Across these exercises, the estimated responses remain stable, indicating that our findings are not driven by the treatment of such borderline episodes. Details are provided in the narrative companion paper (Den Besten and Känzig, 2026).

Figure E.6: Controlling for Unified Government and Partisan Control

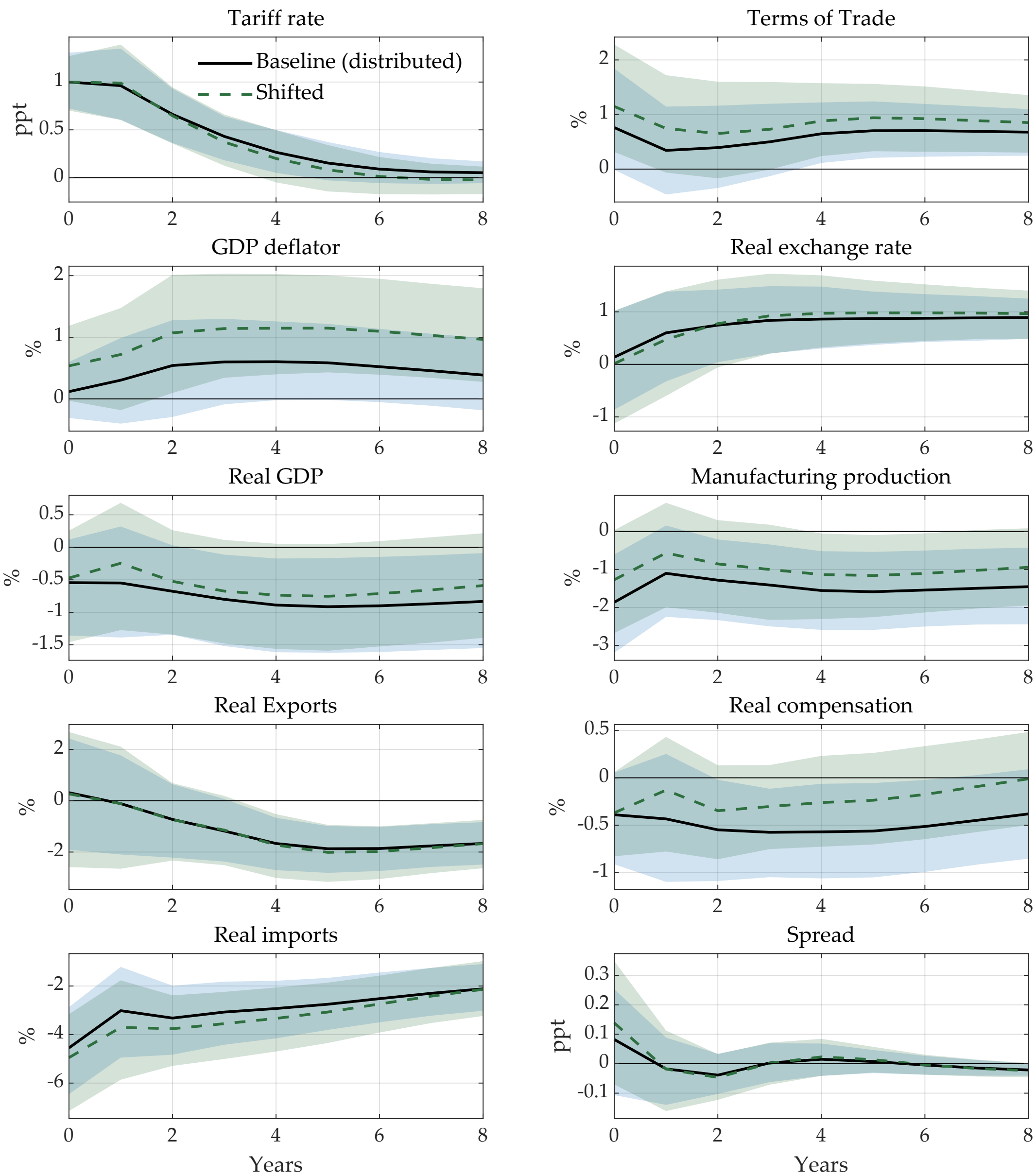


Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated based on the external-instrument VAR model (3). Solid line and shaded areas: baseline response with 90 and 95 percent confidence bands based on moving-block bootstrap. Dashed and dotted lines: point estimates from specifications that include, respectively, a single indicator for unified government and separate indicators for unified Democratic and Republican control.

E.2.2. Timing of the Shocks

As discussed in Section 3.2, we construct our instrument by proportionally allocating each narrative event dummy to the two relevant calendar years, using the share of days remaining in the year after the implementation date as the weight. This approach accounts for the fact that a tariff change implemented in October (e.g., the McKinley Tariff) contributes less to the measured shock in the implementation year than one implemented in January (e.g., the Tariff of 1870). An alternative approach would be to apply a midpoint rule: if a tariff takes effect in the first half of the year, assign it entirely to the current calendar year; if it takes effect in the second half of the year, assign it to the following year. Romer and Romer (2010) follow this rule for their quarterly series.

Figure E.7: Distributed vs. Shifted Narrative Shock



Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated based on the external-instrument VAR model (3). Lag order: 2. Solid black line: baseline responses. Dashed green line: responses using the shifted narrative tariff shock as external instrument. Light shaded areas: 90 percent confidence bands based on moving-block bootstrap.

Figure E.7 compares the results obtained when using the shifted narrative tariff shock constructed with the midpoint rule to our baseline, distributed narrative shock. The two instruments produce similar responses, suggesting that our results are not sensitive to the precise timing assumption used to allocate tariff changes across calendar years.

E.2.3. Size of the Shocks

The key motivation behind our qualitative narrative shock series is twofold. First, while the timing of a tariff event is plausibly exogenous, the magnitude of the tariff change may be endogenous. Second, accurately measuring the size of legislative tariff changes is inherently challenging. Boer and Lütkepohl (2021) show that qualitative sign proxies can provide equally or even more precise identification of impact effects when reliable quantitative measures of shock size are unavailable. Accordingly, we use narrative sign dummies as our baseline instrument, capturing the exogenous timing and direction of tariff changes.

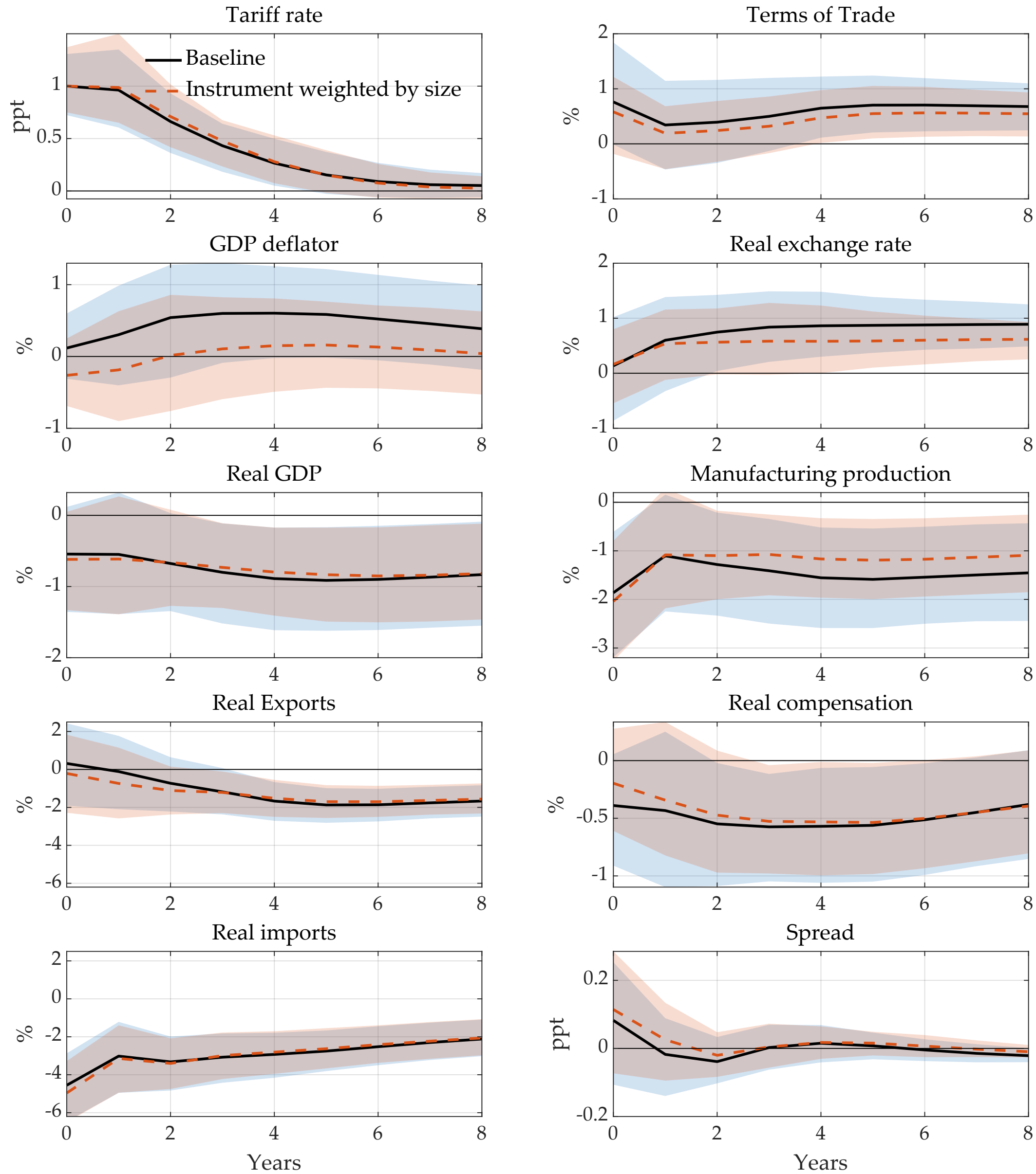
However, relying solely on the sign of an event may provide limited variation and could raise concerns about instrument strength. While the first-stage F-statistic does not indicate a weak instrument problem, we assess whether incorporating quantitative information can improve power.

Measuring the size of a given tariff shock is challenging, as changes in import volumes and, in particular, import prices affect the average tariff rate contemporaneously, making it an imperfect proxy for the magnitude of tariff changes. Ideally, one would hold the composition of imports and their prices fixed and use product-level tariff rates to construct the policy-driven change in tariff revenue for each reform. While such a measure is feasible with more recent data, the necessary product-level information is not readily available for the earlier part of our sample.

In the following, we explore a few different approaches to proxy for the size of a given tariff shock. Our first approach simply weighs the distributed narrative series by the corresponding change in the average tariff rate. Specifically, for each reform occurring at time t, we measure its size as the change in the tariff rate on dutiable imports from $t-1$ to $t+1$. It is important to note that this measure reflects not only the statutory reform itself but also incorporates behavioral responses in imports. For example, if a tariff increase reduces import demand and puts downward pressure on import prices, the ad valorem equivalent of specific duties falls mechanically, thereby dampening the measured size of the tariff change (Irwin, 1998).

Figure E.8 shows the results. Overall, the responses turn out to be similar. The main

Figure E.8: Incorporating the Size of Tariff Changes



Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated based on the external-instrument VAR model (3) using sign-based and size-weighted narrative tariff shocks. Lag order: 2. Solid black line: baseline point estimate. Dashed orange line: point estimate when weighting the instrument by the change in the tariff rate. Light shaded areas: 90 percent confidence bands based on moving-block bootstrap.

difference lies in the deflator response, which is somewhat weaker when using the size-weighted instrument. This likely reflects additional endogeneity introduced by the size-weighted instrument: because tariffs and import prices are inversely related, see for example Irwin (1998), price declines mechanically reduce the ad valorem equivalent of specific duties, dampening the measured size of tariff hikes and weakening the estimated pass-through to prices. This mechanism is consistent with the evidence presented in Figure D.2.

As an alternative size measure, we use estimates by Irwin (1998). To separate import price effects from legislative tariff changes in the average tariff rate, Irwin (1998) regresses the average tariff rate on dutiable imports on the log of the import price index and dummy variables representing tariff legislation and negotiated trade agreements. Accounting for the role of import prices, the coefficients on the tariff dummies represent the policy-determined level of the average tariff rate during each regime, relative to the baseline period. Accordingly, we can use the difference between the tariff coefficients as an estimate of the magnitude of the policy change in the average tariff rate.

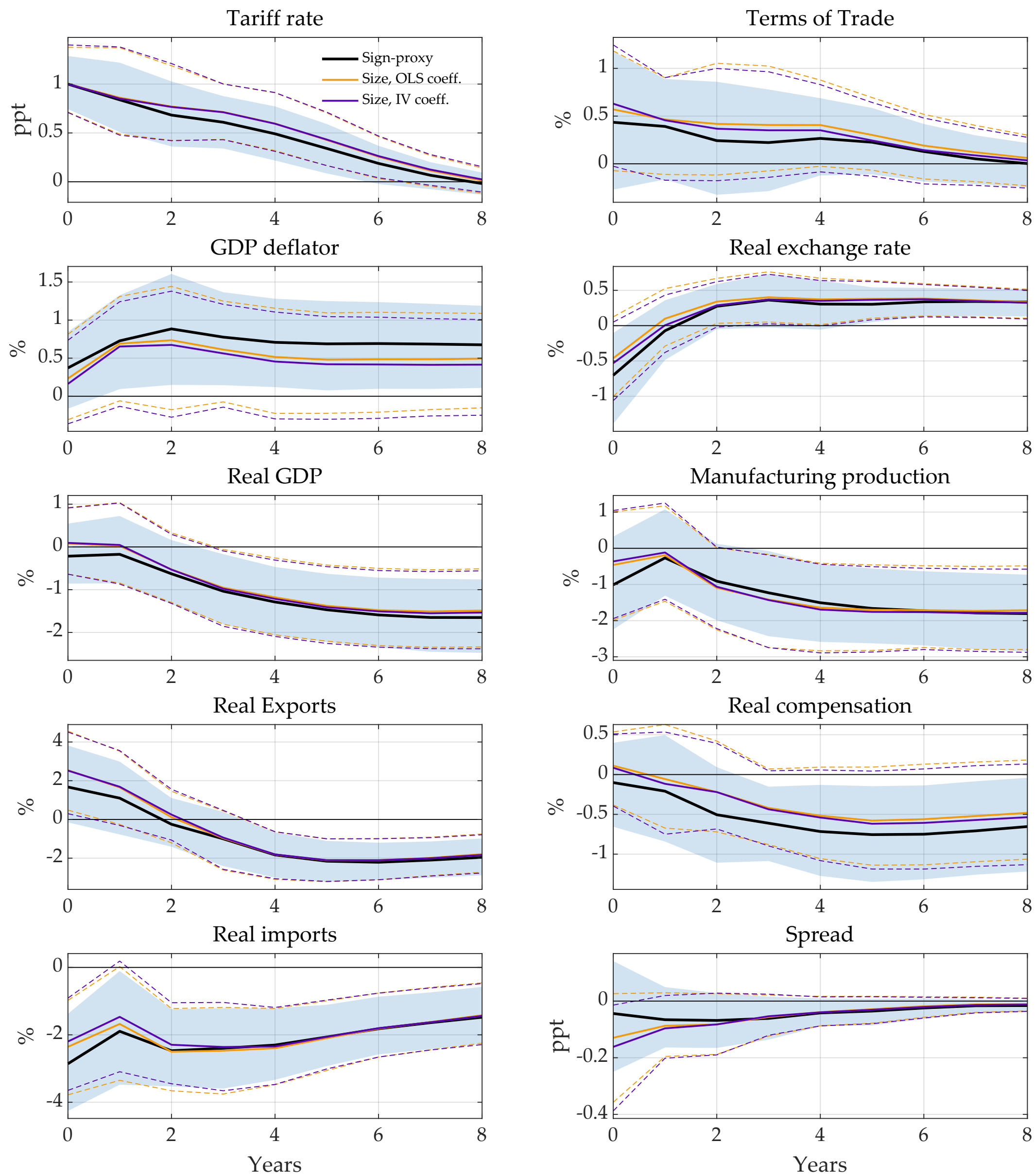
We rely on the coefficients reported by Irwin (1998), rather than re-estimating the specification over our full sample. The reason is that, after the first GATT round, which is the last event included in Irwin's sample, treating our tariff events as discrete regime changes becomes less appropriate. Tariff policy changes increasingly took the form of numerous smaller liberalization steps and temporary or targeted measures, such as Nixon's surcharge on dutiable imports and Ford's fee on imported oil. These actions do not constitute major, broad-based regime changes comparable to the acts of congressional tariff legislation that Irwin models.

We estimate our external-instrument VAR over the period 1865–1967, matching the sample used by Irwin (1998), and construct the instrument from the intersection of Irwin's tariff events with our set of exogenous episodes. Figure E.9 shows that the responses obtained using the narrative sign instrument remain broadly similar to our baseline results, with minor differences reflecting the shorter sample period. Importantly, the results using the size-weighted instrument are very similar as well, although the deflator response is again somewhat weaker. This holds both when using the OLS and IV coefficients in Irwin (1998) (Columns 2 and 3 of Table 2, respectively).

While the size-weighted instrument delivers higher first-stage F-statistics – close to 30 when using the OLS coefficients and close to 40 when using the IV coefficients – this increase in first-stage strength translates into only modest gains in precision, with confidence bands that are comparable to those obtained using the sign-based instrument. At the same time, quantitative measures of tariff changes are more likely to be affected by contemporaneous movements in imports and prices, raising concerns about residual endogeneity.

Taken together, these results suggest that incorporating quantitative information provides limited additional identifying power, while potentially introducing additional endogeneity concerns. This supports our use of narrative sign shocks as the baseline instrument, as they rely on weaker and more plausibly satisfied identifying assumptions while delivering comparable inference.

Figure E.9: Incorporating the Size of Tariff Changes Based on Irwin (1998)

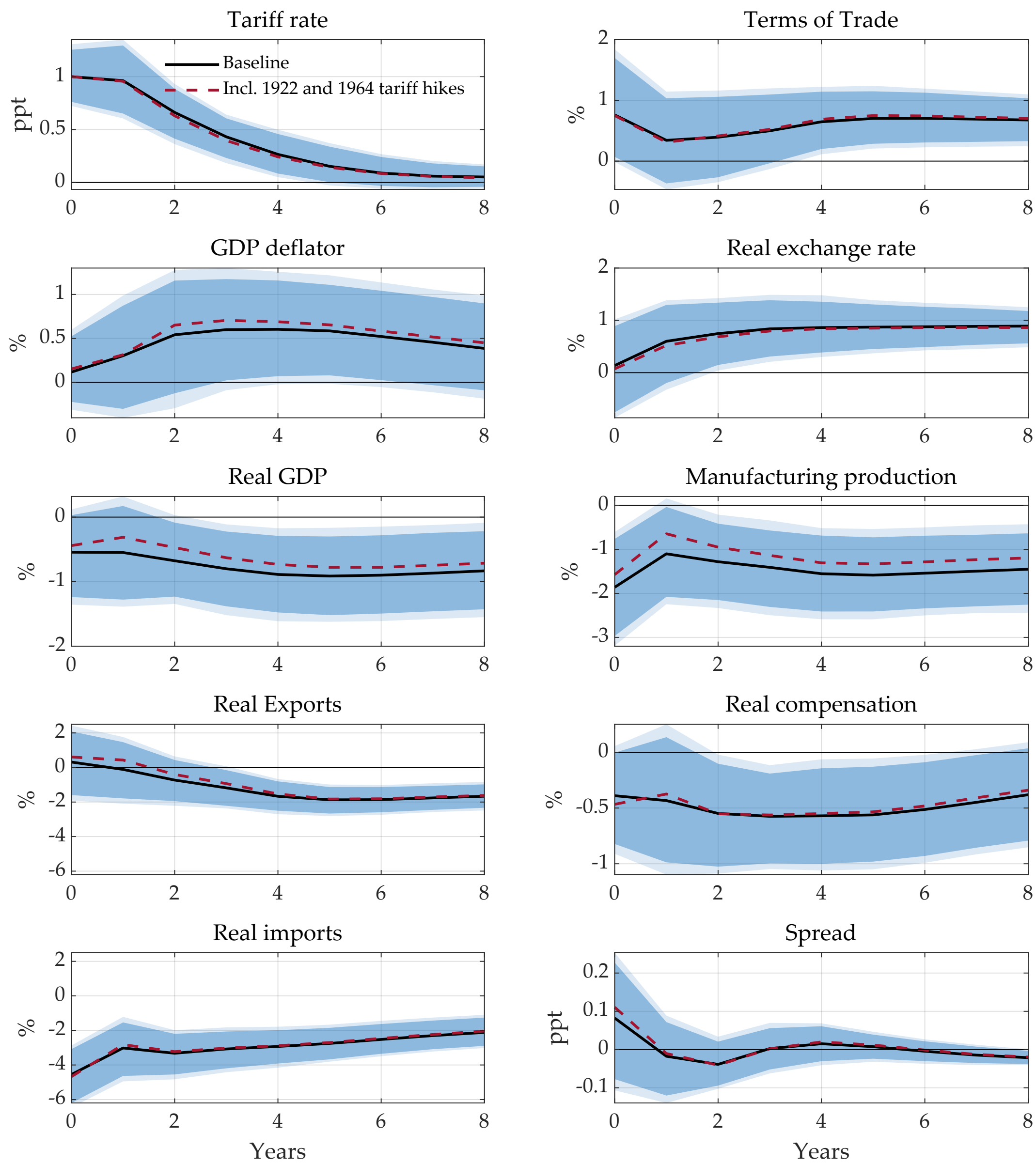


Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated based on the external-instrument VAR model (3) using sign-based and size-weighted narrative tariff shocks. Sample: 1865–1967. Lag order: 2. Solid black line: baseline point estimate. Dashed lines: point estimate when weighting the instrument by the size of the tariff change, based on the coefficients reported by Irwin (1998). The orange line reflects the size adjustment using OLS estimates and the purple line for IV estimates. Light shaded areas: 90 percent confidence bands based on moving-block bootstrap.

E.2.4. Expanding the Number of Positive Shocks

A potential concern is the asymmetry in the number of positive and negative shocks. In our specification, we include fifteen exogenous tariff reductions but only five exogenous tariff increases. This imbalance could matter if the macroeconomic effects of tariff changes are asymmetric by direction.

Figure E.10: Excluding Tariff Hikes



Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated based on the external-instrument VAR model (3), including selected events to the instrument. Solid line and shaded areas: baseline response with 90 and 95 percent confidence bands based on moving-block bootstrap. Dashed line: point estimates when including two tariff hikes to the instrument.

While the small number of positive shocks does not allow us to separately identify the effects of tariff increases and decreases, we assess the robustness of our results to a more balanced specification. Specifically, we include two additional tariff hikes: the Fordney–McCumber Tariff and the 1964 Chicken Tax. Recall, we exclude the Fordney-McCumber Tariff from our baseline specification because the distributed value of the endogenous Emergency Tariff overlaps with the start of the Fordney-McCumber Tariff in 1922. We do not include the Chicken Tax in our baseline specification because this measure applied to a more limited set of product categories.[2]

Figure E.10 shows that our results are robust to a more balanced specification of positive and negative tariff shocks.

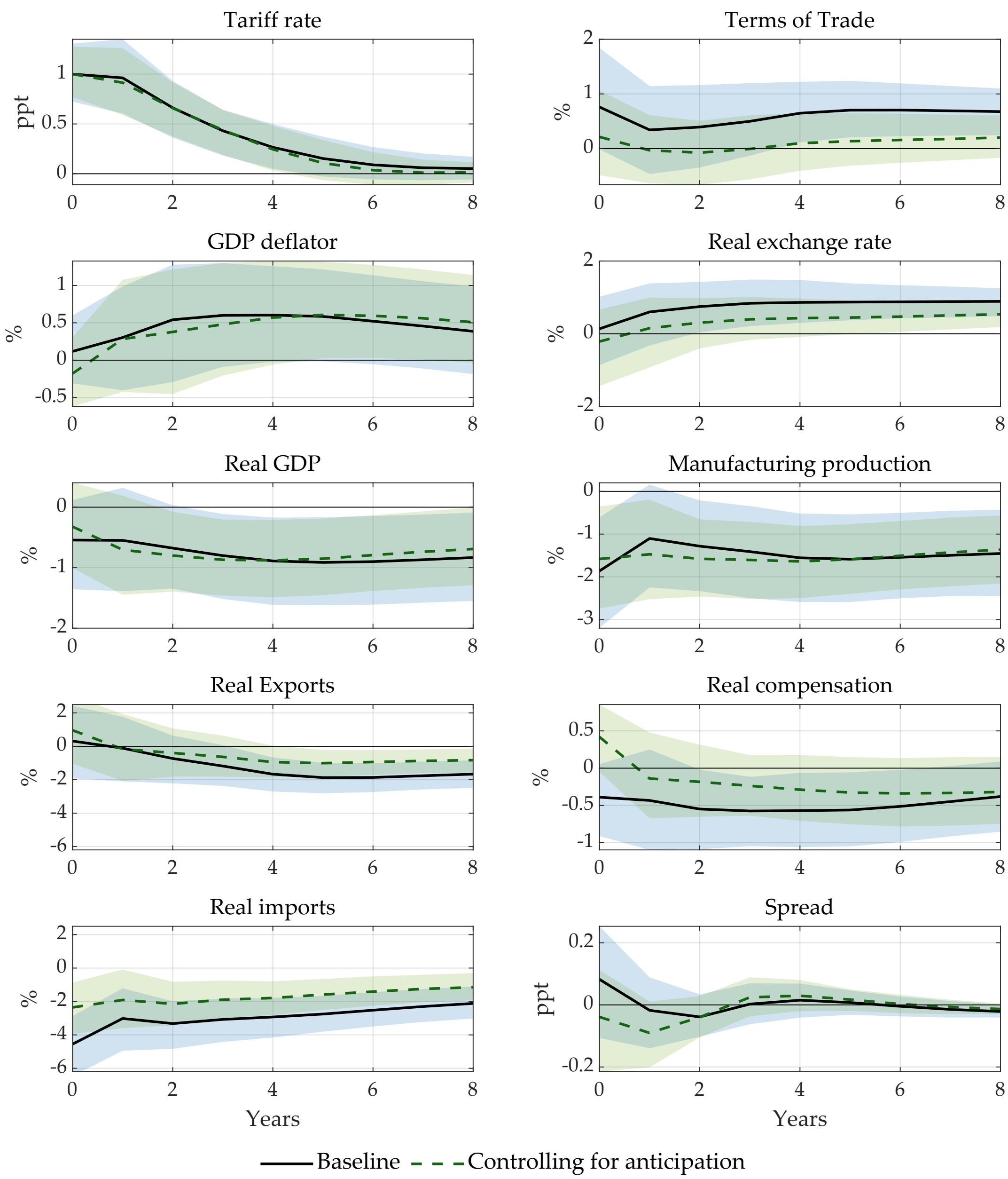
E.3. Accounting for Anticipation Effects

A further concern is that some tariff changes may have been anticipated by firms and households prior to their implementation. If economic agents adjust behavior in advance – for instance by front-loading imports or altering production and pricing decisions – impulse responses centered on the implementation date may fail to capture these preemptive adjustments. As a result, estimated responses could understate or mischaracterize the true effects of tariff shocks. To assess the importance of such anticipation effects, we restrict attention to tariff episodes with minimal delays between legislative approval and implementation.

Figure E.11 shows the impulse responses after excluding tariff episodes with more than a three-month gap between the signing of the legislation and its implementation. The responses are very similar to those in the full sample, suggesting that anticipation effects are unlikely to affect our results.

[2] In an earlier version of the paper, we considered including the Chicken Tax as an exogenous tariff episode. It was implemented in response to European tariffs on American poultry. Following failed negotiations with the EEC, its primary motivation was to exert pressure for policy reversal, reflecting political and strategic considerations. We exclude it from the baseline specification because it is less broad-based than the other tariff changes we analyze

Figure E.11: Controlling for Anticipation Effects



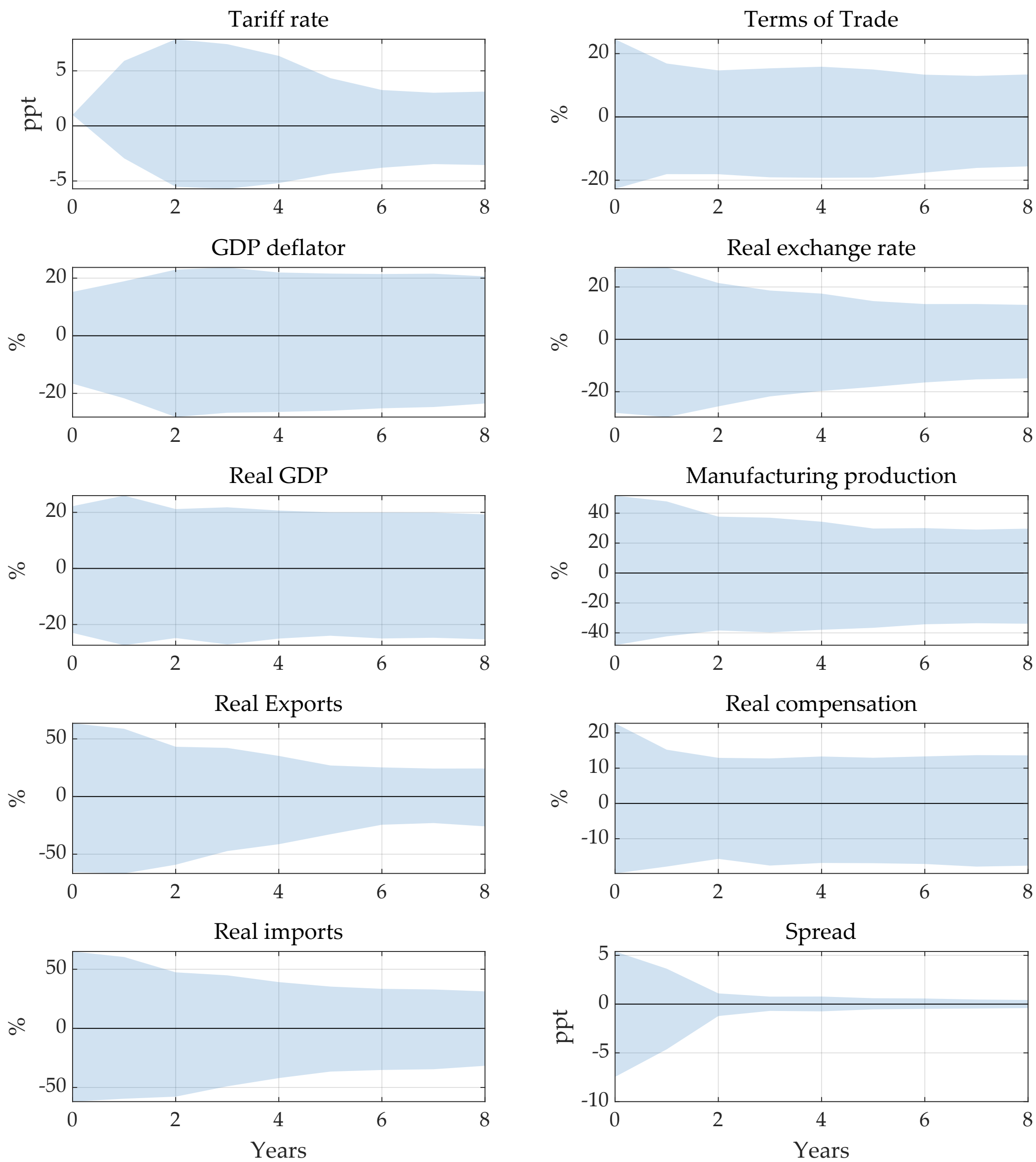
Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated based on the VAR model (3), excluding events with long signing-implementation gaps (over three months) from the instrument. Solid and dashed lines: point estimate. Shaded areas: 95 percent confidence bands based on moving-block bootstrap.

E.4. Placebo Exercise

To further corroborate our identification strategy, we conduct a placebo exercise. We assign random non-event dates to each of the 21 exogenous events in our baseline sample

(1866–2024), such that the number of tariff increases and decreases remains unchanged. For each placebo event date, we reconstruct the distributed proxy by allocating the shock proportionally across the calendar year: we multiply the share of days remaining in the year by the event sign and assign the remainder to the following year. Using this placebo series as an external instrument in our baseline VAR, we estimate impulse responses and repeat this procedure 1,000 times.

Figure E.12: Placebo Responses

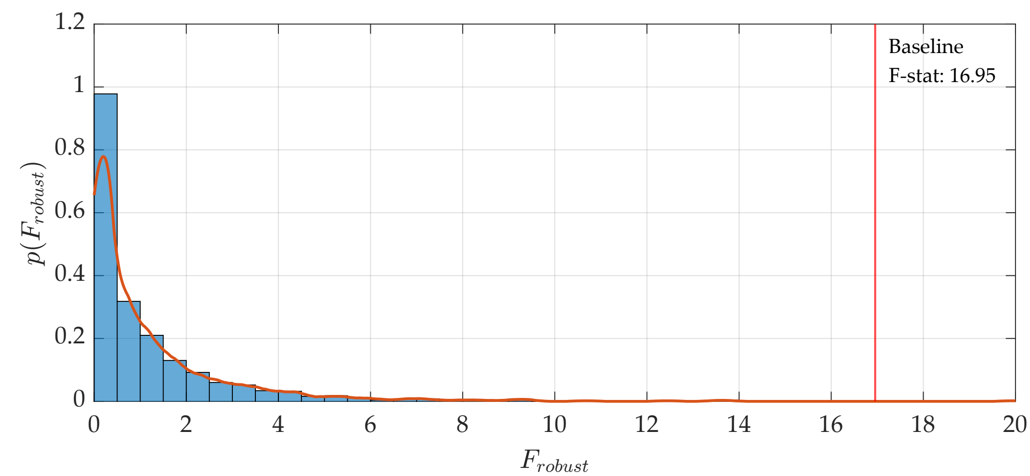


Notes: Shaded areas represent 5th and 95th percentiles of the placebo responses.

Figure E.12 shows the 5th and 95th percentiles of the placebo responses. Exploiting random non-event dates yields impulse responses that do not resemble the systematic macroeconomic effects observed for the actual tariff shocks. The responses do not exhibit any systematic dynamics, and the wide dispersion reflects the fact that the external instrument has little explanatory power for the tariff rate innovation that it instruments.

We can also see this by examining the first stage. Random placebo dates capture a mix of shocks unrelated to tariff policy and therefore should not produce a systematic response in the average tariff rate. As a result, the placebo series should be weak instruments. Figure E.13 confirms this: the distribution of first-stage F-statistics suggests that the placebo instruments have very low explanatory power for the residual of the average tariff rate. Most values cluster near zero, and virtually all mass lies below the conventional threshold of 10, in contrast with the much stronger first-stage statistic obtained for the actual tariff shocks.

Figure E.13: F-statistics of Placebos

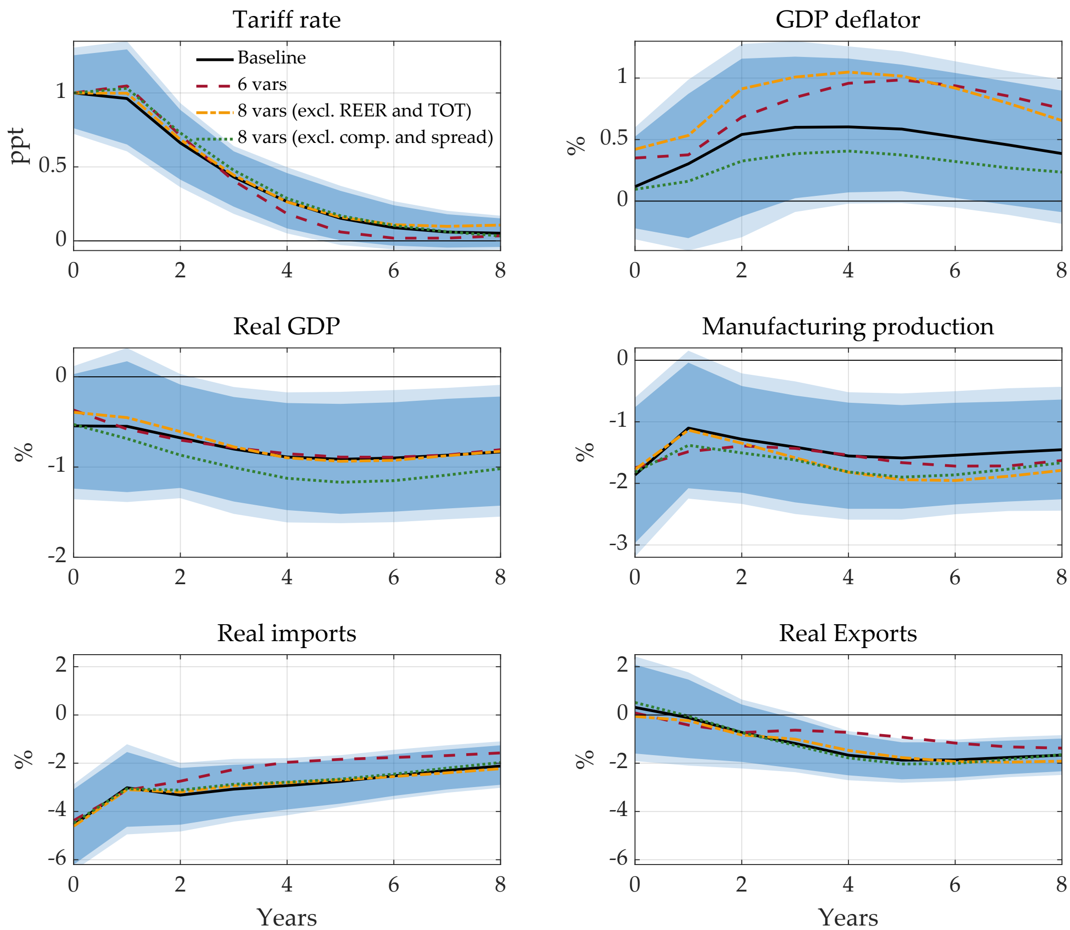


Notes: The distribution is based on the robust first-stage F-statistics obtained from 1,000 placebo instruments. The red vertical line indicates the value of our baseline first-stage F-statistic.

E.5. Specification Choices

In the following section, we examine the sensitivity of the results with respect to the model specification. The baseline VAR includes 10 variables, which is relatively large given our baseline sample. As a robustness test, we use a 6-variable model, excluding the real exchange rate, terms of trade, real compensation of manufacturing production workers, and the interest rate spread. The red dashed line in Figure E.14 shows the point estimates of this 6-variable VAR, which remain consistent with those from the baseline specification. Alternatively, we also estimate 8-variable models. Specifically, we exclude either the real exchange rate and terms of trade (dashed yellow line), or real compensation and the interest rate spread (dashed green line) in Figure E.14. The results remain robust to both the size of the VAR and the specific variables excluded.

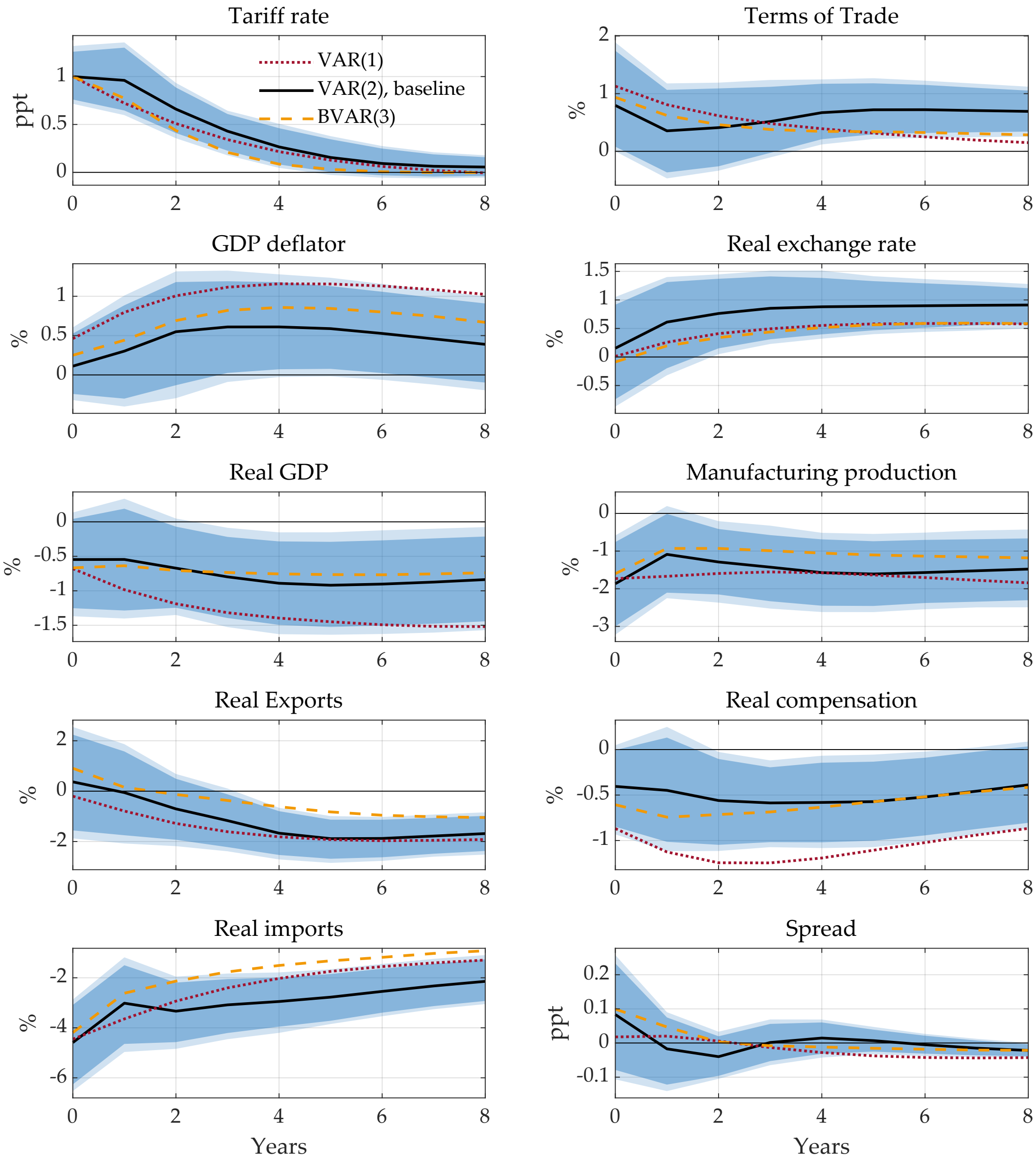
Figure E.14: Responses from Smaller VAR



Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated based on different external-instrument VAR models. Lag order: 2. Solid line and shaded areas: baseline responses with 90 and 95 percent confidence bands based on moving-block bootstrap. Dashed and dotted lines: point estimates based on smaller specifications.

We also assess robustness to alternative lag structures. Figure E.15 compares the baseline VAR(2) specification with a VAR(1) and a Bayesian VAR with three lags estimated using a Minnesota prior with tightness 0.3 and a decay of 2. The responses are highly similar across specifications, with only minor differences in timing and smoothness. In particular, the decline in real GDP is comparable across models, and the overall dynamics of trade, prices, and financial variables remain stable. The BVAR results, which mitigate overfitting concerns in the higher-dimensional specification, closely align with the baseline estimates.

Figure E.15: Sensitivity to Lag Orders

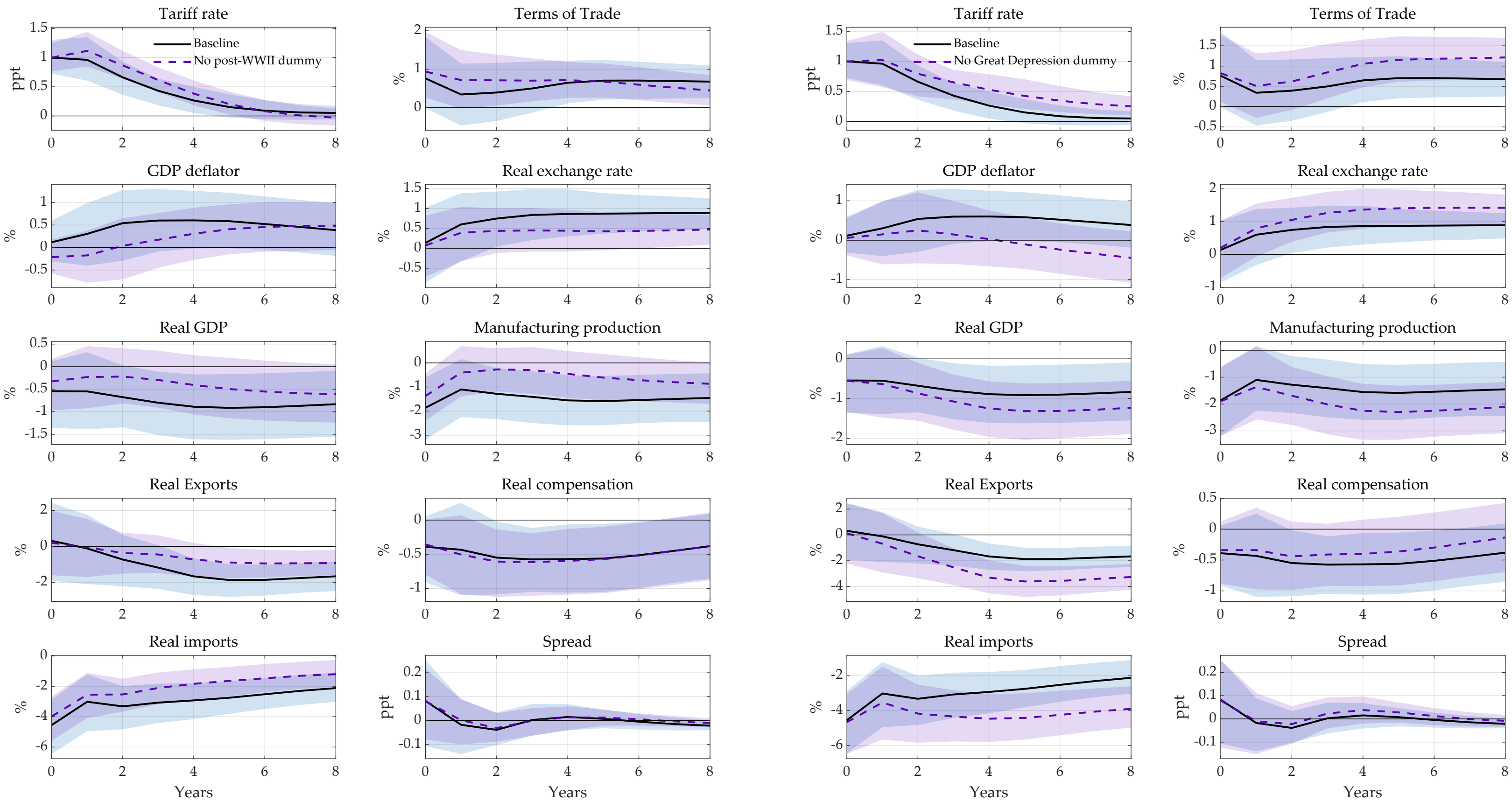


Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated based on the external-instrument VAR model (3) with varying lag orders. Solid black line and shaded areas: baseline responses with 90 and 95 percent confidence bands based on moving-block bootstrap. Dotted and dashed lines: responses with a lag order of 1 and 3, respectively.

Next, we examine the robustness of the results with respect to the deterministic variables included. In our baseline specification, we include a constant, dummies for the two World Wars and the Great Depression, and a post–World War II dummy. These controls account for exceptional historical episodes characterized by large disruptions to economic activity and trade, as well as for the structural shift to persistently lower tariff levels in the postwar period.

Figure E.16: Sensitivity with Respect to Deterministics I

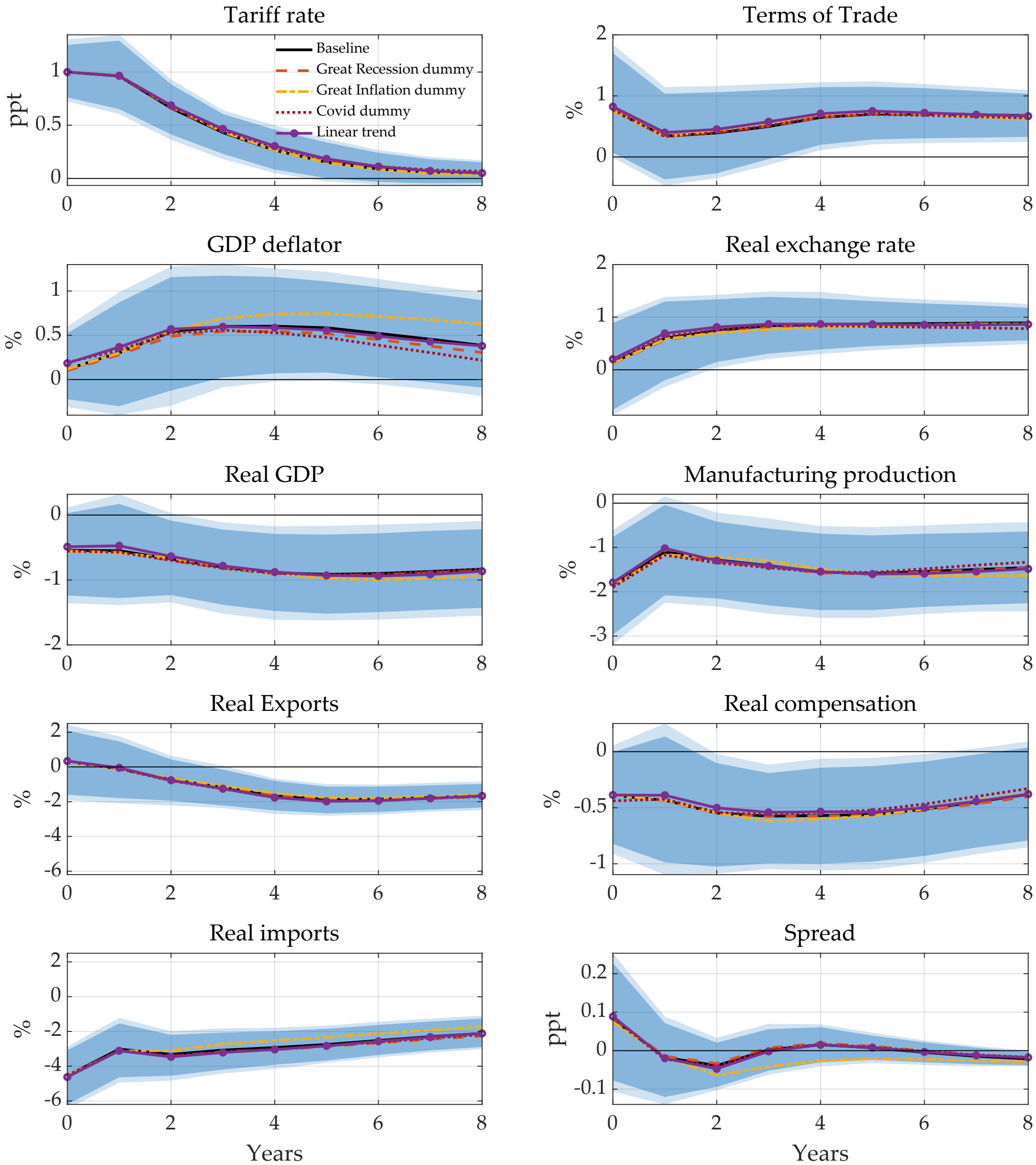
(a) Excluding the Post–World War II Dummy (b) Excluding the Great Depression Dummy



Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated based on the VAR model (3) varying the inclusion of deterministic controls. Panel a: responses when excluding post-WWII dummy. Panel b: responses when excluding Great Depression dummy. Lag order: 2. Solid and dashed lines: point estimates. Light shaded areas: 90 and 95 percent confidence bands based on moving-block bootstrap.

In Figure E.16, we show that our results are robust to the inclusion of the Great Depression or the post-World War II dummy. One slight difference is that when excluding the Great Depression dummy, the deflator response becomes somewhat weaker. This may be because the Great Depression dummy absorbs part of the deflationary dynamics of the early 1930s, which are otherwise attributed to the tariff shock.

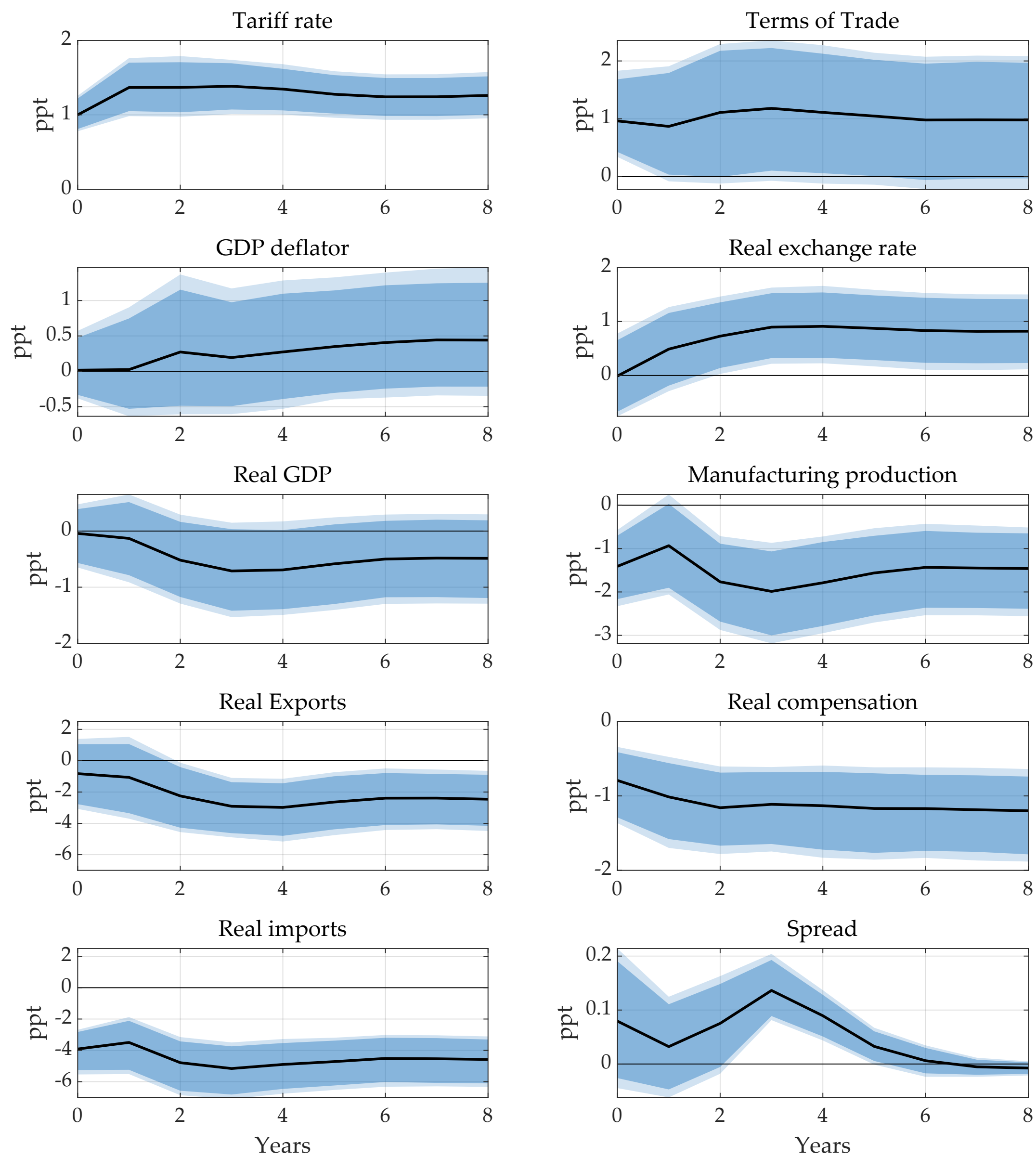
Figure E.17: Sensitivity to Deterministics II



Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated based on the VAR model (3) when controlling for additional deterministic variables. Solid black line and shaded areas: baseline responses and 90 and 95 percent confidence bands based on moving-block bootstrap. Dashed and dotted lines: responses when including additional deterministic controls.

Figure E.17 presents the responses when controlling for additional deterministics. The results are virtually unchanged when controlling for a Great Recession dummy, a Great Inflation dummy, a Covid dummy or including a linear trend. Taken together, these results demonstrate that our estimated impacts are not driven by the particular choice of deterministic terms.

Figure E.18: Specification in Differences



First stage regression: robust F-statistic: 28.13, R^2: 8.59%

Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, based on a VAR estimated in differences with the narrative tariff shock series used as an instrument. Lag order: 2. Solid line: point estimate. Dark and light shaded areas: 90 and 95 percent confidence bands based on moving-block bootstrap.

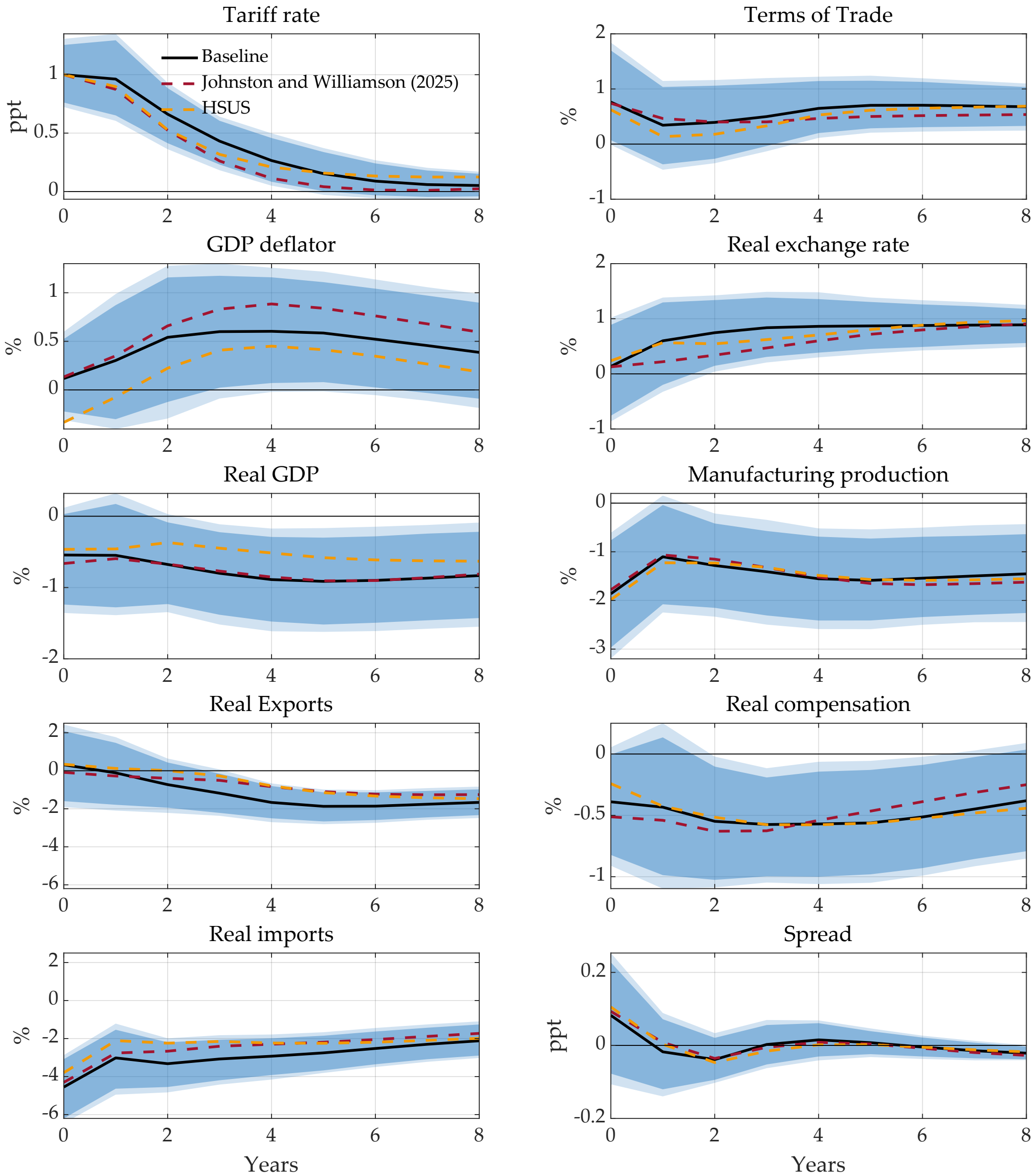
Finally, we estimate the VAR model in differences instead of log-levels. Figure E.18 shows that the responses remain qualitatively similar.

E.6. Alternative Variables and Sources

Historical data are subject to measurement challenges. In this appendix, we show that our results are robust to alternative data and measurement choices.

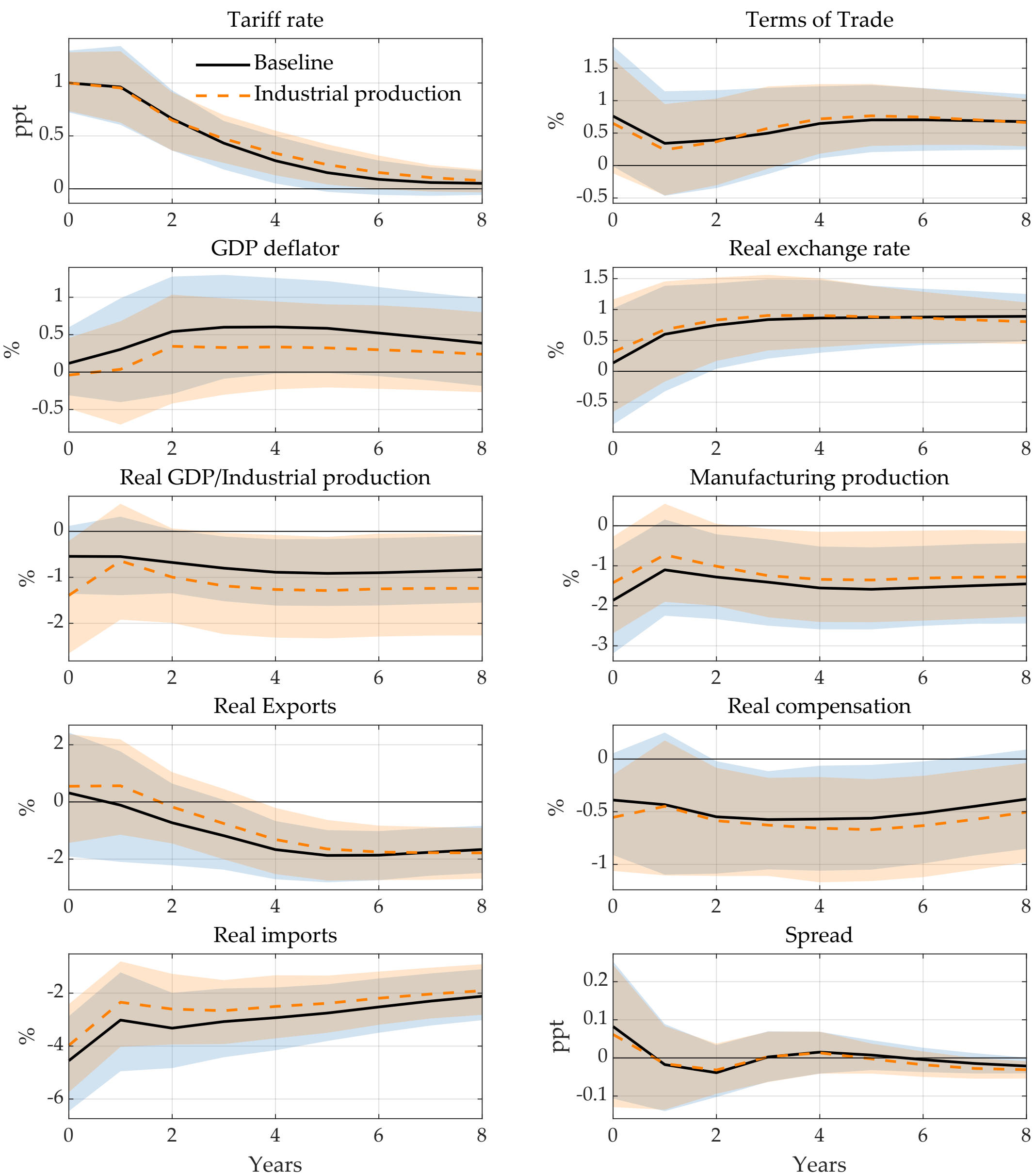
Specifically, we re-estimate the VAR using different series for GDP and the GDP deflator (Figure E.19), replace GDP with industrial production (Figure E.20), and deflate exports and imports using their respective price indices rather than the GDP deflator (Figure E.21). Across all of these alternative constructions of real activity and trade flows, the qualitative and quantitative responses to a tariff shock remain stable, reinforcing the robustness of our main results.

Figure E.19: Responses using Alternative Measures of Output and Prices



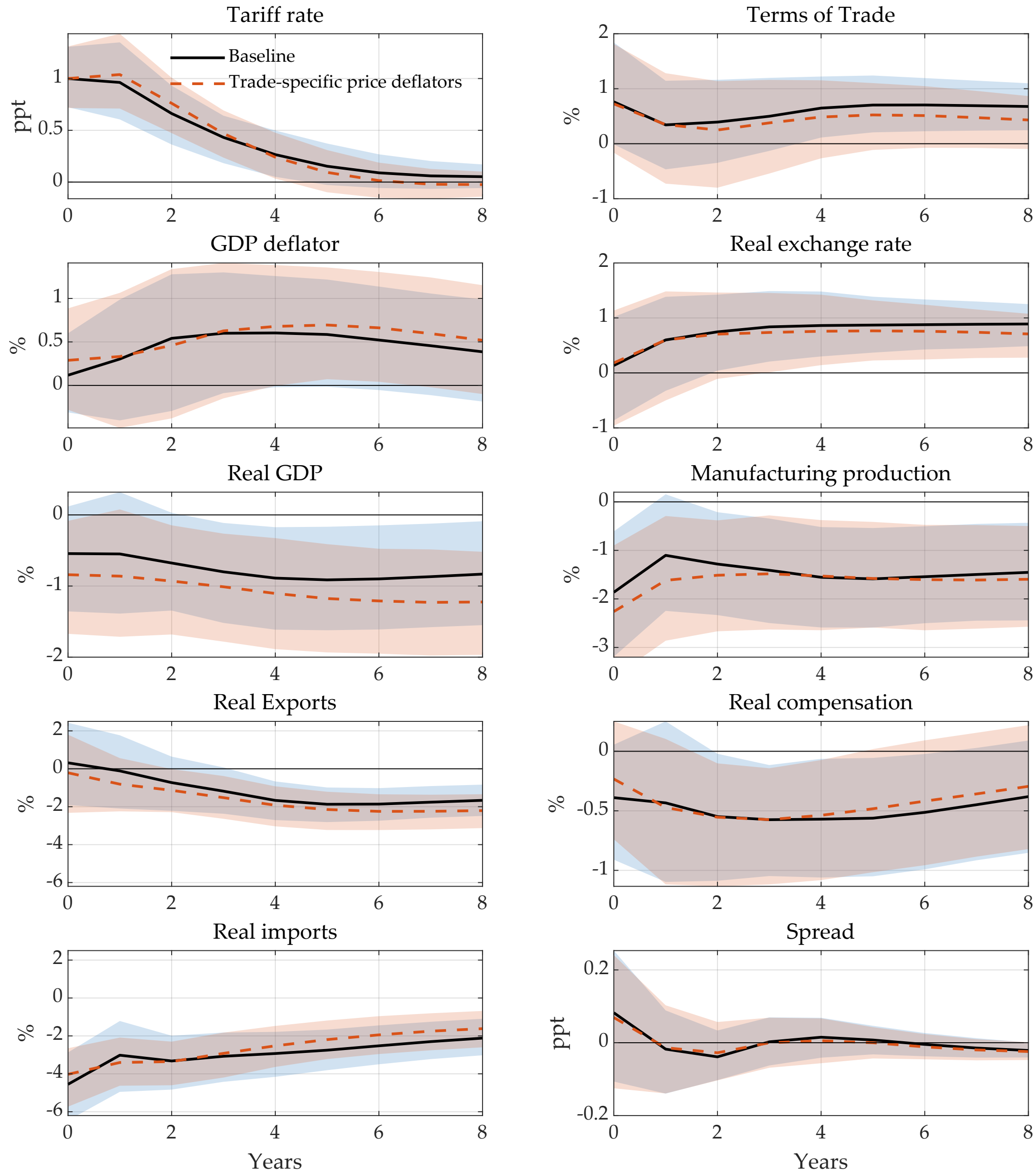
Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated based on the external-instrument VAR model (3) using alternative GDP measures. Lag order: 2. Solid black line and dashed areas: baseline responses with 90 and 95 percent confidence bands based on moving-block bootstrap. Dashed red line: point estimate when using real GDP and GDP deflator data from Johnston and Williamson (2025). Dashed yellow line: point estimate when using real GDP and GDP deflator data from the *Historical Statistics of the United States*.

Figure E.20: Baseline Responses vs. Industrial Production Specification



Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated based on the external-instrument VAR model (3) using alternative activity indicators. Lag order: 2. Solid black line: baseline responses. Dashed orange line: responses when using industrial production instead of real GDP. Light shaded areas: 90 percent confidence bands based on moving-block bootstrap.

Figure E.21: Responses Using Trade-Specific Price Deflators



Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated based on the external-instrument VAR model (3) using different deflators for the trade variables. Lag order: 2. Solid black line: baseline responses. Dashed red line: point estimates when deflating exports and imports by their respective price indices. Light shaded areas: 90 percent confidence bands based on moving-block bootstrap.

References Appendix

Antolin-Diaz, Juan and Paolo Surico (2025). "The long-run effects of government spending". *American Economic Review* 115.7, pp. 2376–2413.

Baker, Scott R., Nicholas Bloom, and Steven J. Davis (2016). "Measuring Economic Policy Uncertainty". *The Quarterly Journal of Economics* 131.4, pp. 1593–1636.

Balke, Nathan S. and Robert J. Gordon (1986). *The American Business Cycle: Continuity and Change*. Ed. by Robert J. Gordon. Studies in Business Cycles. Chicago, IL; London, UK: University of Chicago Press.

Barnichon, Régis and Aayush Singh (2025). *What Is a Tariff Shock? Insights from 150 years of Tariff Policy*. Tech. rep.

Boer, Lukas and Helmut Lütkepohl (2021). "Qualitative versus quantitative external information for proxy vector autoregressive analysis". *Journal of Economic Dynamics and Control* 127, p. 104118.

Bordo, Michael D. (2006). "U.S. monetary gold stock: 1879–1971". *Historical Statistics of the United States, Earliest Times to the Present: Millennial Edition*. Ed. by Susan B. Carter, Scott Sigmund Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch, and Gavin Wright. New York: Cambridge University Press.

Carter, Susan B., Scott Sigmund Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch, and Gavin Wright (2006). "International Trade and Exchange Rates". *Historical Statistics of the United States: Earliest Times to the Present, Millennial Edition*. Cambridge: Cambridge University Press. Chap. Ca.

Chow, Gregory C. and An-loh Lin (1971). "Best Linear Unbiased Interpolation, Distribution, and Extrapolation of Time Series by Related Series". *The Review of Economics and Statistics* 53.4, pp. 372–375.

Davis, Joseph H. (2004). "An annual index of US industrial production, 1790–1915". *The Quarterly Journal of Economics* 119.4, pp. 1177–1215.

Den Besten, Tamar and Diego R. Känzig (2026). "Tariff Changes in the U.S. from 1840–2024: A Narrative Analysis".

Fabricant, Solomon (1940). *The Output of Manufacturing Industries, 1899–1937*. New York: National Bureau of Economic Research.

Fouquin, Michel and Jules Hugot (2016). "Two centuries of bilateral trade and gravity data: 1827-2014". *Vniversitas Económica*, pp. 1–39.

Franconi, Alessandro and Lukas Hack (2025). *Import Tariffs and the Systematic Response of Monetary Policy*. SSRN working paper.

Frickey, Edwin (1947). *Production in the United States, 1860–1914*. Harvard Economic Studies. Cambridge, MA: Harvard University Press, p. 54.

Friedman, Milton and Anna J. Schwartz (1963). *A Monetary History of the United States, 1867–1960*. Studies in Business Cycles. Princeton, NJ: Princeton University Press.

Furceri, Davide, Swarnali A. Hannan, Jonathan D. Ostry, and Andrew K. Rose (2022). "The macroeconomy after tariffs". *The World Bank Economic Review* 36.2, pp. 361–381.

Hanes, Christopher (1998). "Consistent Wholesale Price Series for the United States, 1860–1990". *Business Cycles since 1820: New International Perspectives from Historical Evidence*. Ed. by Trevor J. O. Dick. Cheltenham, UK: Edward Elgar, p. 205.

Imlah, Albert H. (1958). *Economic elements in the Pax Britannica: Studies in British foreign trade in the nineteenth century*. Harvard University Press.

Irwin, Douglas A. (1998). "Changes in US tariffs: the role of import prices and commercial policies". *The American Economic Review* 88.4, pp. 1015–1026.

— (2006). "Exports and Imports, by Broad End-Use Class: 1923–2001". *Historical Statistics of the United States, Earliest Times to the Present: Millennial Edition*. Ed. by Susan B. Carter, Scott Sigmund Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch, and Gavin Wright. New York: Cambridge University Press, Table Ee506–Ee520.

— (2017). *Clashing over commerce: A history of U.S. trade policy*. University of Chicago Press.

Jacks, David S. (2019). *Real commodity prices, 1850–present*. https://davidjacks.org/from-boom-to-bust/.

Jaramillo-Vera, Jaime A. (2025). *tempdisagg: A Python Library for Temporal Disaggregation of Time Series*.

Johnston, Louis and Samuel H. Williamson (2025). *What Was the U.S. GDP Then?* MeasuringWorth.

Klau, Marc and San Sau Fung (2006). "The new BIS effective exchange rate indices". *BIS Quarterly Review*.

Lévy-Leboyer, Maurice and François Bourguignon (1985). *L'Economie Française au XIXe siècle: Analyse macro-économique*. Paris: Economica.

Lipsey, Robert E. (1963). *Price and Quantity Trends in the Foreign Trade of the United States*. Princeton, NJ: Princeton University Press.

Miranda-Agrippino, Silvia and Giovanni Ricco (2021). "The transmission of monetary policy shocks". *American Economic Journal: Macroeconomics* 13.3, pp. 74–107.

Miron, Jeffrey A. and Christina D. Romer (1990). "A new monthly index of industrial production, 1884–1940". *The Journal of Economic History* 50.2, pp. 321–337.

Mitchell, B. R., ed. (2013). *International Historical Statistics*. Springer book series.

Müller, Karsten, Chenzi Xu, Mohamed Lehbib, and Ziliang Chen (2025). *The Global Macro Database: A New International Macroeconomic Dataset*. Working Paper 33714. National Bureau of Economic Research.

North, Douglass C. (1961). *The Economic Growth of the United States, 1790–1860*. Englewood Cliffs, NJ: Prentice Hall.

Officer, Lawrence H. and Samuel H. Williamson (2025). *Annual Wages in the United States, 1774–Present*. MeasuringWorth.

Plagborg-Møller, Mikkel and Christian K. Wolf (2019). "Local projections and VARs estimate the same impulse responses".

Ramey, Valerie A. and Sarah Zubairy (2018). "Government spending multipliers in good times and in bad: evidence from US historical data". *Journal of Political Economy* 126.2, pp. 850–901.

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, pp. 763–801.

— (2023). "Presidential address: Does monetary policy matter? The narrative approach after 35 years". *American Economic Review* 113.6, pp. 1395–1423.

Schwert, G. William (1990). "Indexes of US Stock Prices from 1802 to 1987". *Journal of business*, pp. 399–426.

Simon, Matthew (1960). "The United States balance of payments, 1861-1900". *Trends in the American economy in the nineteenth century*, pp. 629–716.

Taussig, Frank W. (1910). *The Tariff History of the United States*. 5th. New York and London: G.P. Putnam's Sons / The Knickerbocker Press.

U.S. Census Bureau (1949). *Historical Statistics of the United States, 1789–1945: A Supplement to the Statistical Abstract of the United States*. Washington, D.C.: U.S. Government Printing Office.

Warren, George F. and Frank A. Pearson (1932). *Wholesale Prices for 213 Years, 1720 to 1932*. Memoir 142. Ithaca, NY: Cornell University Agricultural Experiment Station, pp. 7–10, 84–111.

Wright, Gavin (2006). "Energy consumption, by energy source: 1850–2001". *Historical Statistics of the United States, Earliest Times to the Present: Millennial Edition*. Ed. by Susan B. Carter, Scott Sigmund Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch, and Gavin Wright. New York: Cambridge University Press.