The Macroeconomic Effects of Tariffs: Evidence From U.S. Historical Data*

Tamar den Besten† Diego R. Känzig‡

February, 2026

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.

JEL classification: E30, F13, F14, F41, H20

Keywords: Trade policy, tariff shocks, narrative identification, macroeconomic effects

*We thank Marios Angeletos, Regis Barnichon, Javier Bianchi, Larry Christiano, Marty Eichenbaum, Lukas Hack, Douglas Irwin, Joel Mokyr, Giorgio Primiceri, Matt Rognlie, Kunal Sangani, Ludwig Straub and Chris Wolf for helpful comments and discussions.

†Northwestern University. E-mail: tamardenbesten2029@u.northwestern.edu. Web: www.tamardenbesten.com.

‡Northwestern University, CEPR and NBER. E-mail: dkaenzig@northwestern.edu. Web: www.diegokaenzig.com.

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 mirror 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, average tariff rates embed behavioral responses affecting imports and prices, so changes in these measures often reflect movements in economic conditions rather than the underlying policy action. 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 demand 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 modest 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 indicate that demand effects are also important, particularly through weaker exports and broader macroeconomic contraction. The simultaneous presence of supply-side inflationary pressures and demand-driven disinflation helps 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, demonstrating that the contractionary 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 a tariff shock.

These differences point to stronger demand effects in the modern era, consistent 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 on impact. Both these mechanisms exert downward pressure on domestic prices, relative to the gold standard period, consistent with the disinflationary response observed in the postwar subsample. 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.

Accounting for the endogeneity of tariff policy is central to uncovering these effects. A simple recursive identification that treats tariff movements as exogenous innovations yields markedly different responses of output, prices, and trade flows. 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 a broad set of robustness and sensitivity checks. We assess the role of identifying assumptions, including tests for invertibility and alternative dynamic lag structures; 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 remarkably stable.

Overall, the central message is clear: tariff increases depress economic activity and trade once their indirect and general-equilibrium effects are taken into account. Although tariffs may appear protective in partial equilibrium, the broader economy experiences sizable declines in output, manufacturing activity, and external competitiveness. These effects are transmitted through a mix of 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, reshaping 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 consequences 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 direct evidence on how tariffs shape output, prices, and trade flows, but they face two important 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) and government spending (Ramey and Shapiro, 1998; Ramey, 2011). We port this narrative methodology to tariff policy, constructing a new long-run series of plausibly exogenous tariff changes spanning 1840–2024.

A closely related, contemporaneous paper by Barnichon and Singh (2025) studies the effects of tariff policy on unemployment and inflation using historical data for the United States, the United Kingdom, and France. They present results based on OLS and a narrative instrumental-variable approach, and emphasize demand-side mechanisms, including uncertainty. Our work contributes to this literature by providing a more comprehensive historical account of U.S. tariff policy and by placing greater emphasis on isolating tariff shocks from non-policy variation, large confounding macroeconomic events, and structural breaks – leading to more nuanced conclusions about inflation dynamics.

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 (e.g., Werning, Lorenzoni, and Guerrieri, 2025; Bergin and Corsetti, 2023; Monacelli, 2025; Bianchi and Coulibaly, 2025; Auray, Devereux, and Eyquem, 2025). 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 today 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 (ADTR):

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

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, 1998). 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 shows that this fixed-rate measure is highly correlated with the dutiable tariff rate, concluding that the latter provides a reliable proxy for changes in overall tariff protection when only variable-weight data are available.

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 (1998), 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 D.4 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).

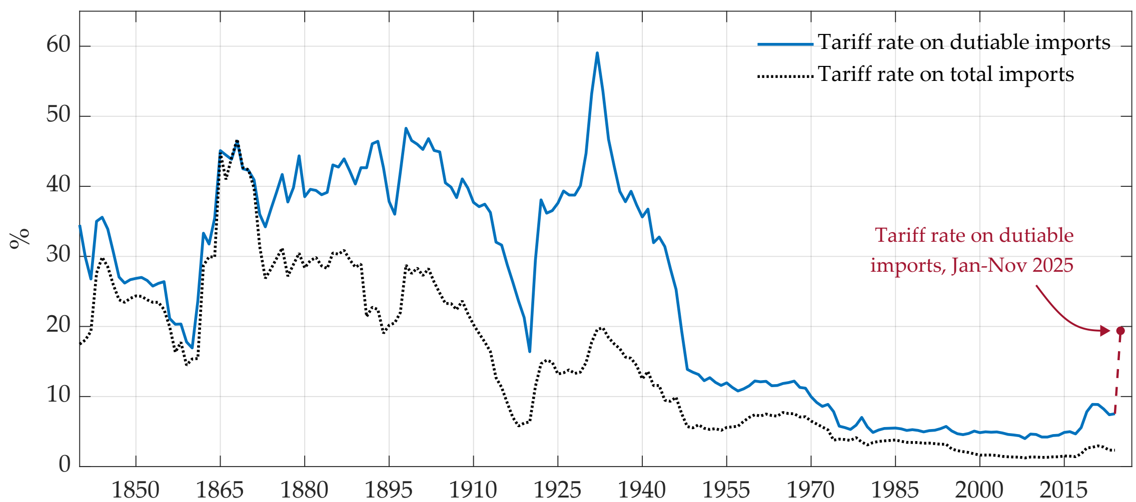
Because quarterly data are unavailable for many variables in the early sample, we follow Balke and Gordon (1986) and apply the Chow–Lin temporal disaggregation method to distribute annual observations across quarters using relevant high-frequency indicators, while preserving annual totals. As in Antolin-Diaz 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 custom duties calculated over imports, from 1840 to 2024 and provides a visual summary of this history.

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 as essential for American industrial development. 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).

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 economic conditions, marked by weak growth and high unemployment throughout much of the 1890s, led to opposite tariff responses depending on political majorities. 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 farm and import-competing interests, the core logic of Smoot–Hawley was still distributive and protectionist, and the bill was introduced before the 1929 stock market crash. 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 today. 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 today. 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 $Z_t \in \{-1, 1\}$ 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 therefore helps isolate the component of policy change that is most likely to reflect an autonomous shift in trade policy.

As the baseline, we thus construct an annual series that records only the timing and sign of each exogenous tariff episode. In robustness exercises, we consider alternative scalings based on changes in the average tariff rate on dutiable and total imports (see Appendix D.1).

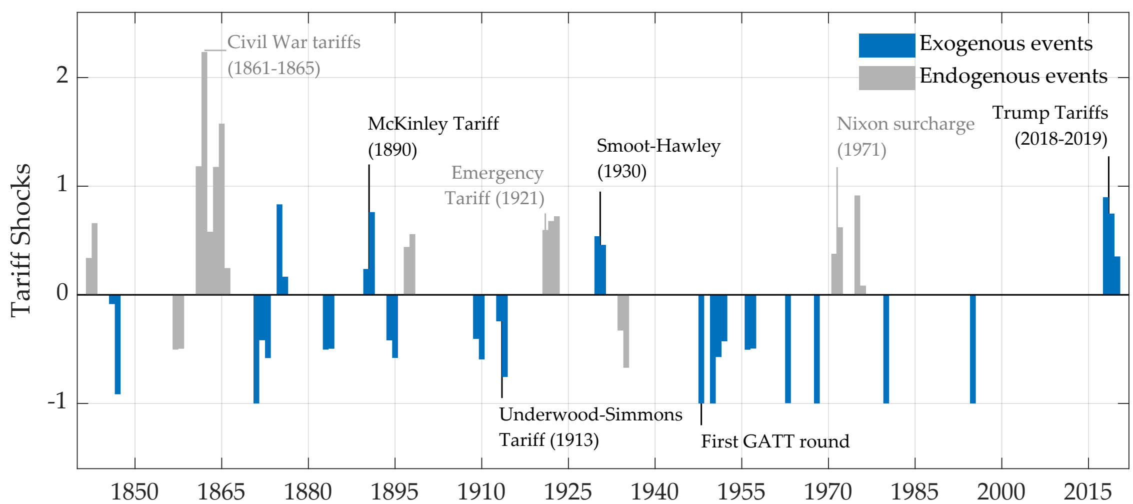
Because tariff reforms are implemented on specific calendar dates rather than at uniform intervals, we distribute each narrative shock across the two adjacent years in proportion to the number of days remaining in the year of implementation. Let d_t denote the number of days left in year t after the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

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

where $Z_t \in \{-1, 1\}$ 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. Tariff increases occur primarily in the earlier part of the sample, whereas the postwar period is characterized predominantly by exogenous tariff cuts.

Figure 2: Narrative Tariff Shock Series



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

Diagnostics. We perform a series of diagnostic checks on the narrative shock series (Ramey, 2016). We find limited evidence that the narrative tariff shock series is serially correlated. We also find no evidence that macroeconomic or financial variables, such as output, exports, imports, the terms of trade, stock prices or government finances, have any power in forecasting the narrative tariff shock series. Table 1 reports the p-values for the corresponding Granger causality tests. For all variables considered, the p-values are far above conventional significance levels, with the joint test having a p-value of 0.92, 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	0.15	0.00
Tariff rate	0.89	0.07
Terms of trade	0.92	0.60
Real exports	0.13	0.10
Real imports	0.18	0.05
Macroeconomic variables		
Real GDP	0.30	0.26
Manufacturing production	0.38	0.09
Real compensation	0.57	0.75
GDP deflator	0.74	0.03
Producer price index	0.45	0.19
Financial variables		
Interest rate spread	0.95	0.06
REER	0.54	0.97
Real energy commodity price index	0.54	0.12
SP500	0.57	0.00
Government finances		
Real government debt	0.71	0.82
Real government revenue	0.78	0.45
Real government expenditure	0.26	0.37
Joint	0.92	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. Number of lags: 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.[3] 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.

Finally, we show that the series is also uncorrelated with other shock measures put forward in the literature, including different measures of fiscal and monetary policy shocks, see Appendix C.

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 *instrument*. 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:

$$\mathbf{y}_t = \mathbf{B}(L)\mathbf{S}\boldsymbol{\varepsilon}_t, \tag{1}$$

where $\boldsymbol{\varepsilon}_t$ is a vector of mutually uncorrelated structural shocks driving the economy, $\mathbf{B}(L) \equiv \mathbf{I} + \mathbf{B}_1 L + \mathbf{B}_2 L^2 + \dots$ is a matrix lag polynomial, and $\mathbf{S}$ is the structural impact matrix.

Assuming that the vector-moving average process (1) is invertible, it admits the following VAR representation:

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

where $\mathbf{u}_t$ is a $n \times 1$ vector of reduced-form innovations with variance-covariance matrix

[3] 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.

$\text{Var}(\mathbf{u}_t) = \mathbf{\Sigma}$ 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, z_t is the narrative tariff shock series. For z_t to be a valid instrument, we need:

$$\mathbb{E}[z_t \varepsilon_{1,t}] = \alpha \neq 0 \tag{3}$$

$$\mathbb{E}[z_t \boldsymbol{\varepsilon}_{2:n,t}] = \mathbf{0}, \tag{4}$$

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 (3) is the relevance requirement and assumption (4) is the exogeneity condition. Under these assumptions, the structural impact vector $\mathbf{s}_1$ is identified up to sign and scale:

$$\mathbf{s}_1 \equiv \boldsymbol{\theta}_0 \propto \frac{\mathbb{E}[z_t \mathbf{u}_t]}{\mathbb{E}[z_t \mathrm{u}_{1,t}]}, \tag{5}$$

provided that $E[z_t \mathrm{u}_{1,t}] \neq 0$. Note that to identify the relative impact effects, we do not need to assume invertibility. Invertibility is only required for the dynamic effects beyond impact.[4]

To facilitate interpretation, we scale the structural impact vector such that the shock corresponds to a 1 percent 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, as proposed by 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

[4]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).

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:

$$\mathbb{E}[z_t \boldsymbol{\varepsilon}_{t+j}] = \mathbf{0}, \quad \text{for } j \neq 0. \tag{6}$$

Together with the exogeneity and relevance requirement, 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, $\bar{\mathbf{s}}_1 = [\text{chol}(\bar{\mathbf{\Sigma}})]_{\cdot,1} / [\text{chol}(\bar{\mathbf{\Sigma}})]_{1,1}$.

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 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:

$$y_{i,t+h} - y_{i,t-1} = \alpha_h^i + \theta_h^i \Delta y_{1,t} + \boldsymbol{\beta}_h^{i\prime} \Delta \mathbf{x}_{t-1} + \nu_{i,t,h}, \tag{7}$$

using z_t as an instrument for $y_{1,t}$. 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, grouped into several blocks. The international and trade block contains the average dutiable tariff rate, imports, exports, the terms of trade, and the real effective exchange rate. The output and production block includes real GDP and manufacturing production. The prices and wages block consists of the GDP deflator and the compensation of production workers, allowing us to track both aggregate price dynamics and labor-cost pressures. Finally, the financial conditions block includes 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

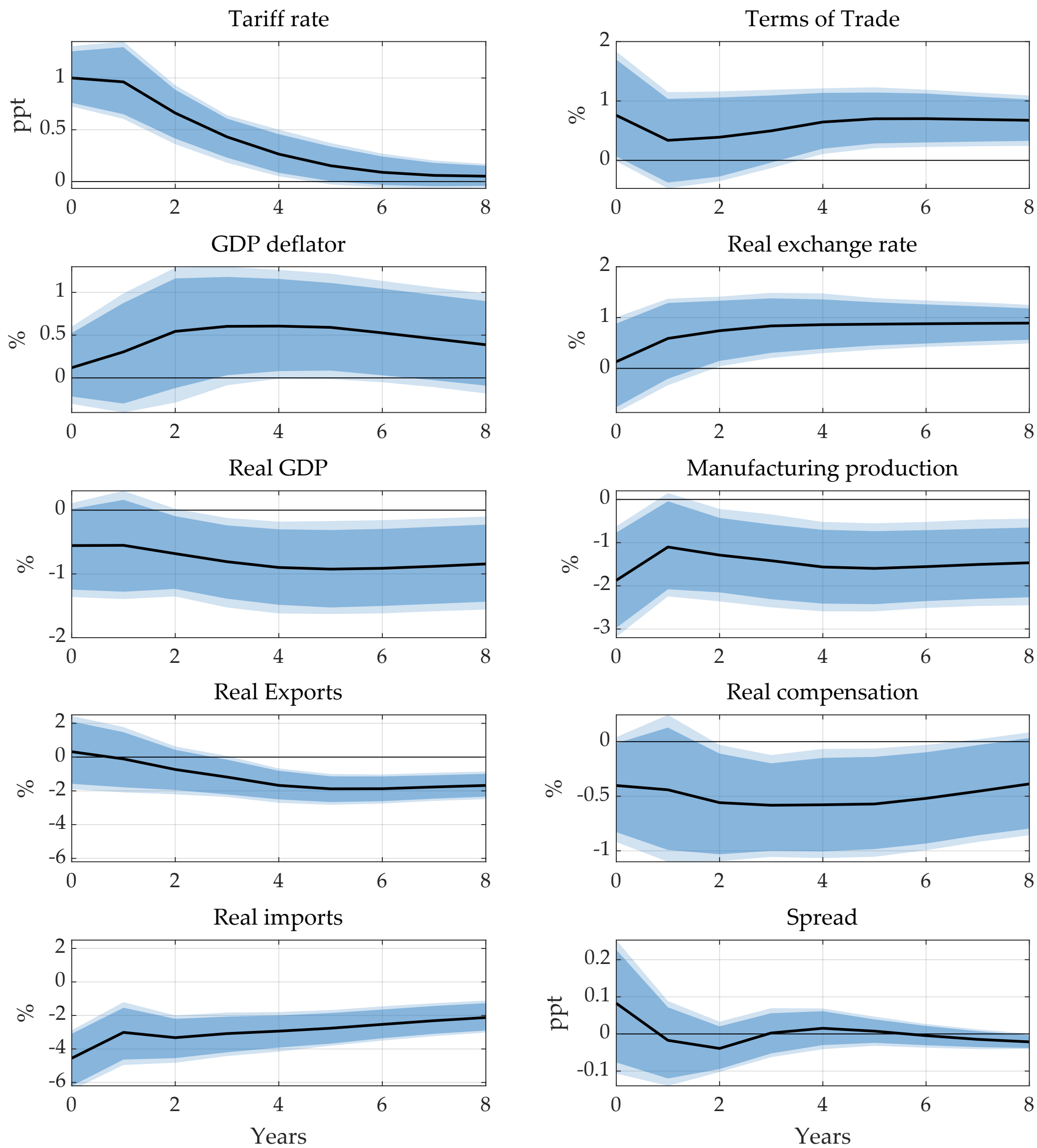
5.1. The Macroeconomic Effects of Tariffs

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 16.96 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 3 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 one percentage point on impact. 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.

Figure 3: The Macroeconomic Effects of a Tariff Shock



First stage regression: robust F-statistic: 16.96, R^2: 6.22%

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 (2) 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.

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 little and the point estimates are 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).

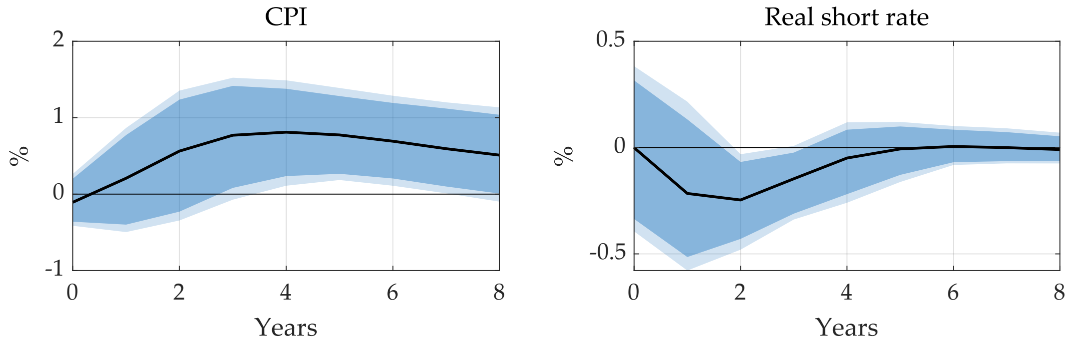
The terms of trade improve following a tariff shock, consistent with standard trade theory. Part of this improvement reflects a valuation effect: the appreciation of the domestic currency mechanically lowers import prices in domestic currency relative to export prices. However, the importance of this 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, 1998). This pattern is consistent with the classic optimal-tariff channel. However, despite this relative price gain, tariff shocks are contractionary overall, underscoring that 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 sustained inflation at the aggregate level. The importance of these demand-side channels is consistent with the mechanisms emphasized by Barnichon and Singh (2025), although in our setting they attenuate rather than overturn the short-run inflationary response.

Figure 4 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 tariffs feeding more directly into consumer-facing prices than into broader aggregate price measures. 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 3. 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).

Figure 4: The Effect on Consumer Prices and Interest Rates



Notes: Impulse responses of consumer prices and the real short-term interest rate 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.

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.

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.

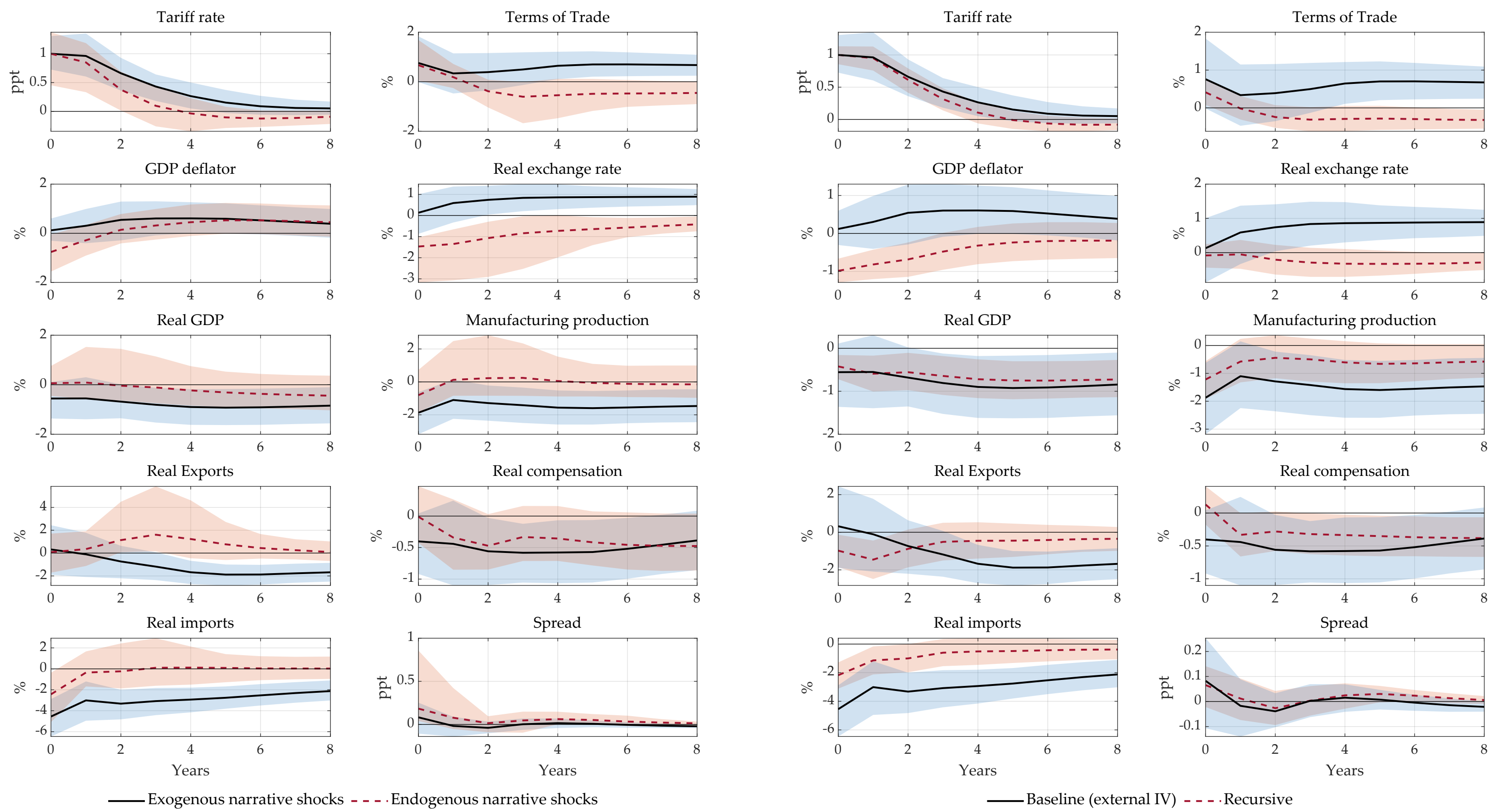
5.2. The Importance of Accounting for Endogeneity in Tariff Changes

A key motivation for our narrative identification strategy is the concern that tariff changes are often endogenous to economic conditions. Historically, many tariff reforms were implemented in response to spending needs or the business cycle, as opposed to motivations exogenous to the economy. In this section, we illustrate the importance of accounting for this endogeneity by comparing our baseline results to two alternative identification

Figure 5: The Role of Accounting for the Endogeneity in Tariff Changes

(a) Exogenous vs. Endogenous Tariff Shocks

(b) SVAR-IV vs. Recursive Identification



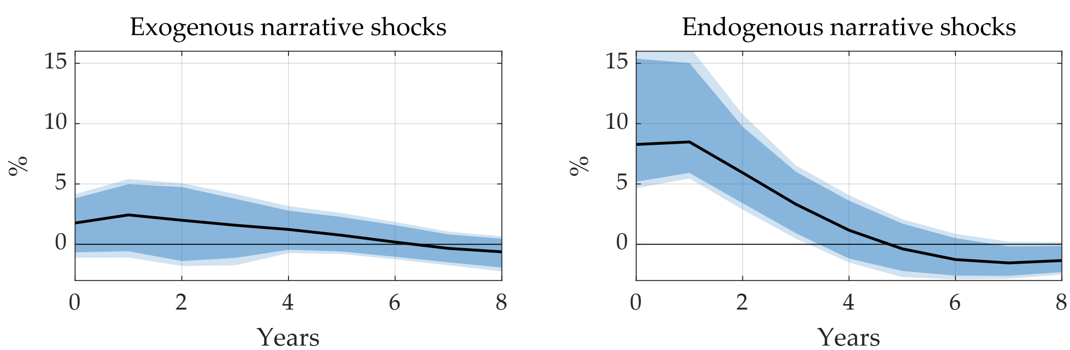
Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact. Panel a: responses using exogenous and endogenous narrative tariff shocks as the instrument, respectively. Panel b: baseline responses against recursive VAR. Lag order: 2. Solid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

strategies that do not account for the endogeneity in tariff changes.

Exogenous versus endogenous tariff events. We begin by contrasting the impulse responses obtained using our baseline instrument – constructed from plausibly exogenous tariff events – with those obtained when the instrument is instead based on tariff events classified as endogenous. The resulting responses are shown in Figure 5a. The responses based on the endogenous narrative shocks differ markedly from our baseline findings. When tariff changes are motivated by endogenous considerations, imports and exports exhibit little systematic response, and output shows no significant decline. The exchange rate tends to depreciate rather than appreciate, 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 interpretation, we show in Figure 6 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 underscore that failing to account for the endogeneity of tariff rates can lead to estimated responses that differ markedly from the true effects of tariff shocks that are exogenous to the state of the economy.

Figure 6: The Effect on Government Spending



Notes: Government spending response to endogenous and exogenous tariff shocks, estimated based on the external instrument VAR model (2). 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.

Comparison with recursive approach. We next compare our external-instrument approach to a standard recursive VAR, with the average tariff rate ordered first. This approach assumes that the tariff rate is predetermined with respect to the macroeconomy, effectively treating all tariff changes as exogenous innovations.

As discussed in Section 3, this is a strong assumption that is unlikely to hold in practice, especially at the annual frequency. Many historical tariff reforms were motivated by factors endogenous to the economy. Moreover, even when underlying policy changes are exogenous, measures of average tariff rates incorporate non-policy-induced variation arising from behavioral responses. As a result, recursively identified tariff shocks conflate endogenous variation with exogenous policy shocks.

Consistent with these identification concerns, the resulting impulse responses differ sharply from our baseline. In the recursive model, the responses of imports and exports are again 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 now falls significantly and persistently. Notably, this pattern – declining output and prices following tariff changes – is in line with recent evidence in Barnichon and Singh (2025), which relies on a similar historical sample. However, our results show that once the endogeneity of tariff changes is properly accounted for, tariff shocks exhibit stagflationary properties (see Appendix D.3 for a more detailed comparison to Barnichon and Singh 2025).

Taken together, these exercises demonstrate that failure to account for the endogeneity of tariff changes can lead to misleading inferences about their macroeconomic effects. Both the use of endogenous tariff events and the assumption that tariff rates are predetermined generate responses that differ not only in magnitude but also in sign from our baseline estimates. This evidence reinforces the importance of a careful narrative classification and instrumental-variable strategy to isolate plausibly exogenous tariff shocks and obtain credible estimates of their causal effects on trade, output, prices, and financial variables.

5.3. The Propagation of Tariff Shocks Over Time

To assess the stability of our results over time, we re-estimate the baseline VAR on a series of sub-samples 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.

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. Consistent with 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. This shift points to a stronger role for demand-side channels in the transmission of tariff shocks in the modern era.

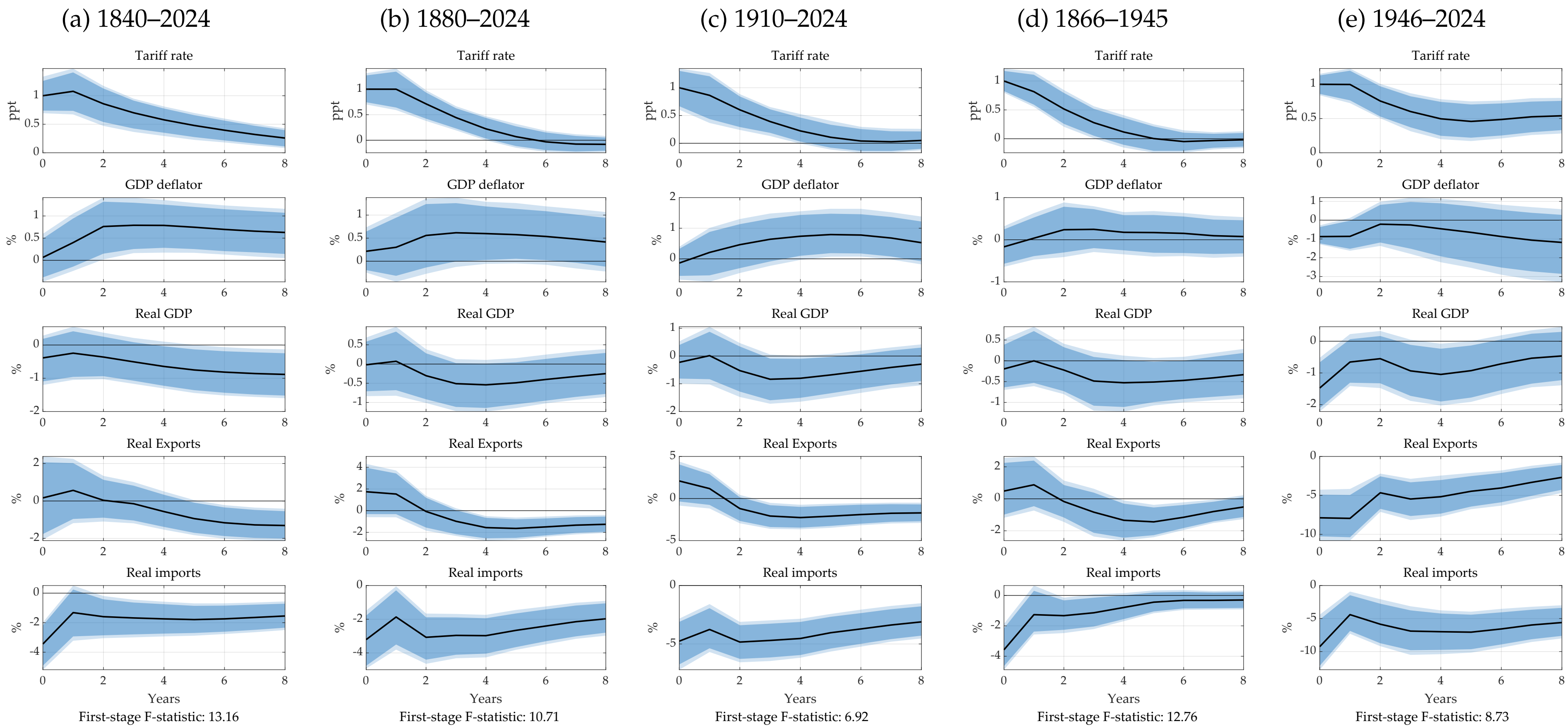
Importantly, when the postwar period is included as part of a longer sample, the estimated responses largely resemble those obtained from prewar data alone. This reflects the fact that much of the identifying variation in tariff shocks – and many of the largest tariff changes – occur prior to World War 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.[5] 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 playing a more prominent role in the postwar period.

5.4. The Role of Domestic and Foreign Policy Responses

A first explanation for the differences in tariff transmission across samples relates to changes in the monetary regime, in particular the transition from gold-backed mone-

[5] In Appendix D.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.

Figure 7: Macroeconomic Effects of Tariff Shocks Across Sub-samples



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 (2), 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.

tary 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.

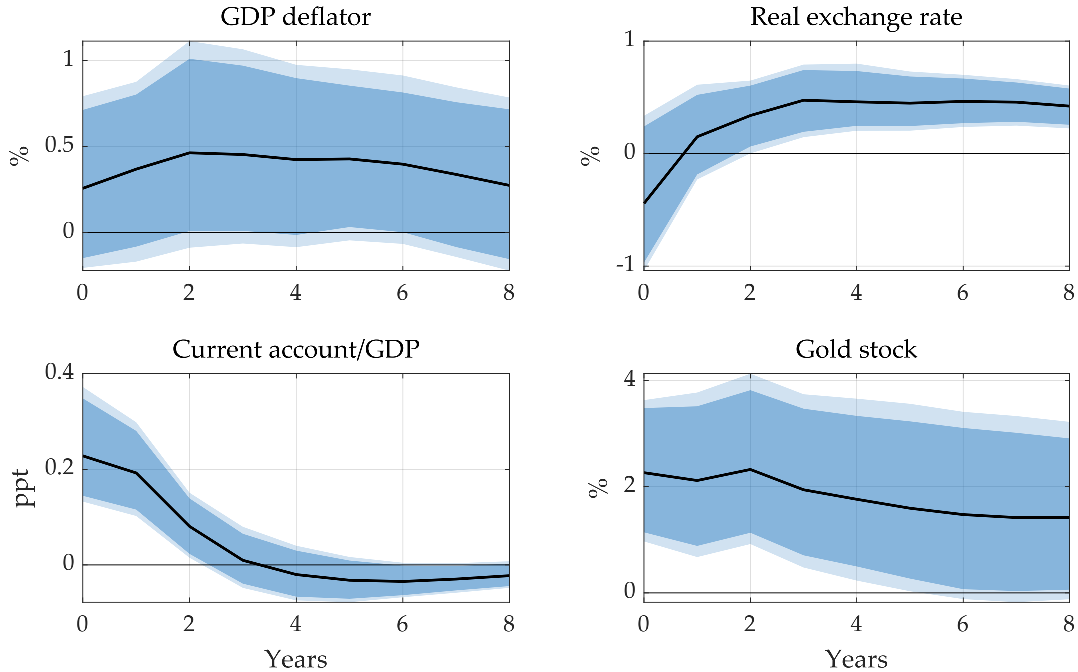
Under the gold standard, nominal exchange rates were effectively fixed, implying that real exchange rate adjustment could occur only through changes in relative prices. Following a tariff increase, imports fall sharply on impact, improving the trade balance. Because the nominal exchange rate cannot appreciate, 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 pattern closely mirrors 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 patterns using a model estimated over the era of gold-backed monetary systems, spanning 1879–1971. Figure 8 presents the corresponding impulse responses for a subset of variables. Consistent with the mechanisms outlined above, a tariff shock generates an inflationary response of the GDP deflator and a gradual real appreciation driven by rising domestic prices. At the same time, the trade balance improves, as reflected in a positive response of the current account. This improvement is accompanied by gold inflows, leading to a persistent increase in the stock of gold. Appendix Figure E.2 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 9, which shows the responses of the real exchange rate, the current account, and the real short-term interest rate in the pre- and post-World War II periods. In the prewar sample, the real exchange rate appreciates only gradually,

Figure 8: Responses for Era of Gold-Backed Monetary Systems, 1879-1971

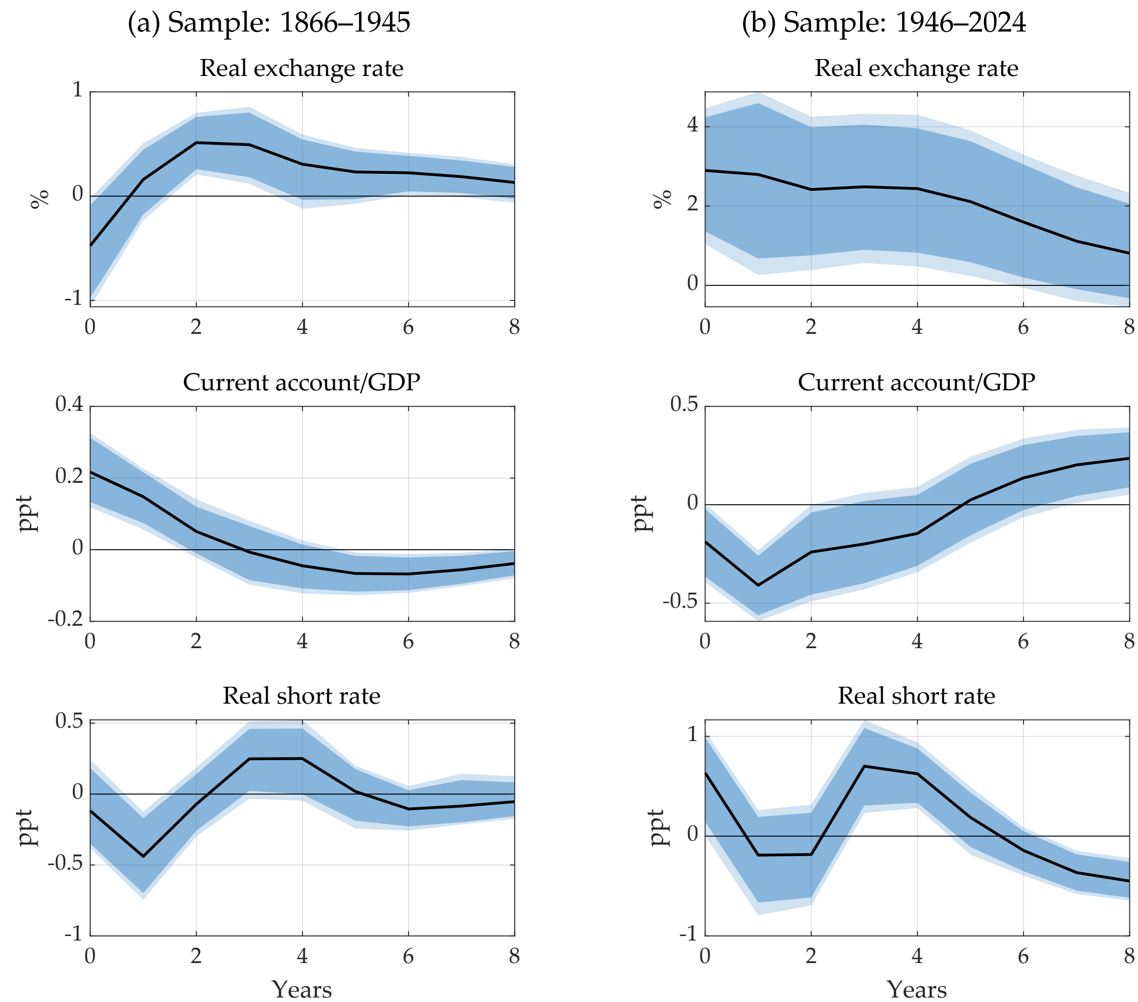


Notes: Impulse responses to a tariff shock over the era of gold-backed monetary systems, normalized to increase the dutiable tariff rate by 1 percentage point on impact. The responses are estimated based on the VAR model (2), including the current account over GDP and the stock of gold in lieu of the 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.

the current account improves, and the real short rate declines, consistent with a more accommodative monetary stance under gold-backed monetary arrangements. In contrast, in the post-World War II period, the real exchange rate appreciates immediately, the current account deteriorates, and the real short rate rises temporarily, indicating a markedly different adjustment mechanism. 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. In the postwar period, tariff policy became increasingly embedded in multilateral negotiations and reciprocal trade agreements, with most policy changes taking the form of coordinated tariff reductions. Against this background, unilateral tariff increases were more likely to be perceived as deviations from an established cooperative framework and therefore to elicit rapid foreign responses, including retaliation, reciprocal tariff adjustments, or shifts in sourcing patterns. This stands in contrast to the prewar era, when tariff changes were more often unilateral and less tightly embedded in a system of reciprocal commitments.

Figure 9: External Adjustment and Monetary Response: 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, estimated based on the VAR model (2). Panel a: pre-World War II sample, 1866–1945; Panel b: post-World War II sample, 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.

As a result, the contractionary effects of tariffs in the postwar period operate more forcefully through exports and aggregate demand. Rapid foreign responses weaken external demand for U.S. goods, reinforcing the effects of exchange rate appreciation and domestic demand spillovers. In this environment, tariff shocks propagate primarily through demand-side channels, generating disinflationary pressures despite higher trade barriers and helping to explain the sharper and more immediate export and price responses observed in the modern era.

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

5.5. Sensitivity Analyses

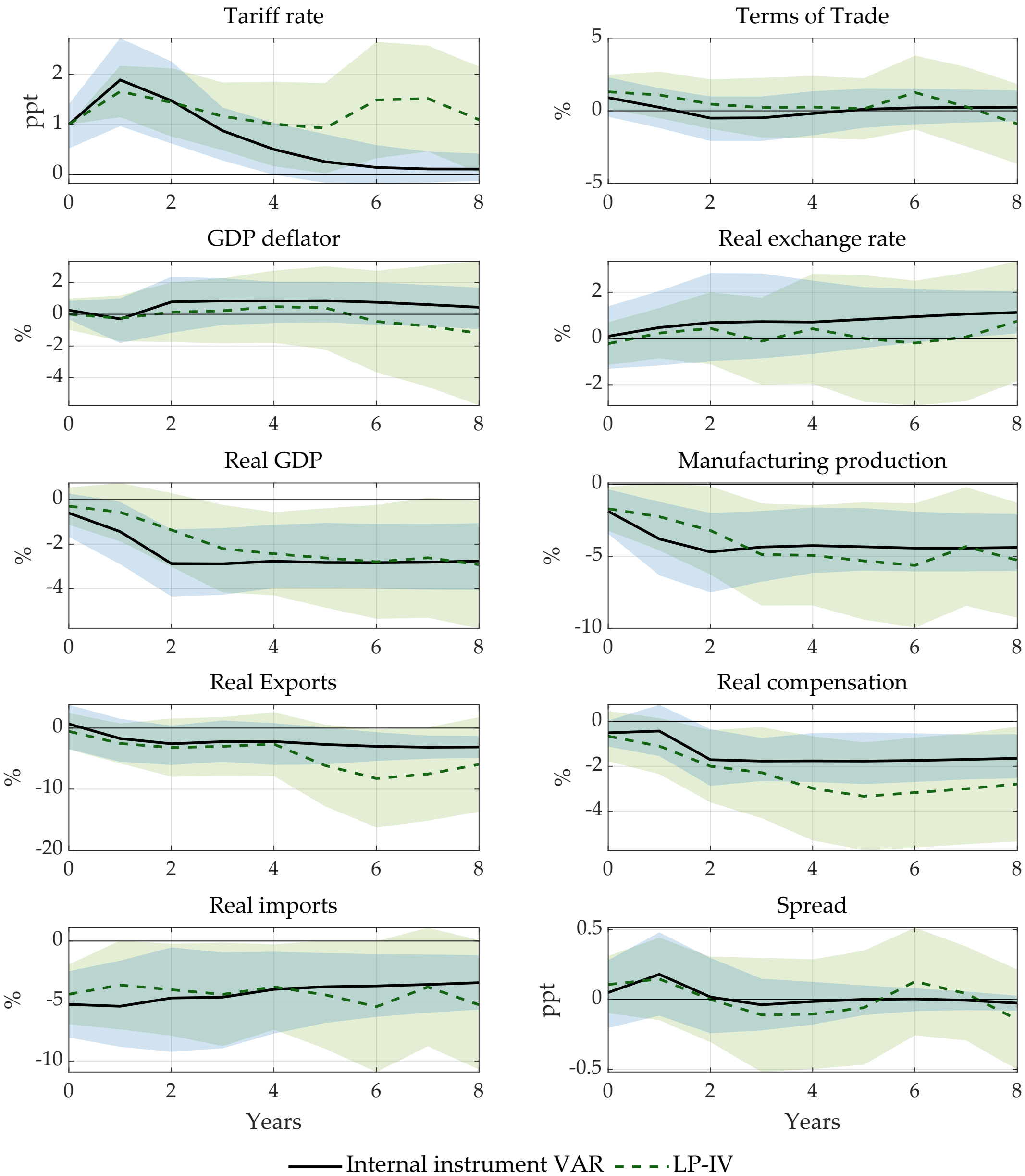
In this section, we perform a comprehensive series of sensitivity checks. Specifically, we assess the importance of the assumptions underlying our VAR approach, including invertibility and the dynamic structure, and investigate the role of potential confounders and anticipation effects.

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 10 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.

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. A natural concern, however, is that even when tariffs were not primarily motivated by fiscal considerations or economic conditions, such factors may have been present in the background during the legislative process, potentially inducing residual correlation with underlying macroeconomic shocks.

Figure 10: Relaxing VAR Assumptions



Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated based on alternative models: the internal instrument VAR and the LP-IV approach. Solid and dashed line: point estimates. Shaded areas: 90 percent confidence bands based on moving-block bootstrap.

To address this concern, we examine the sensitivity of our results to alternative narrative classifications, with particular attention to tariff episodes in which fiscal imbalances or broader economic pressures played a secondary or contextual role rather than a central motivating role. As documented in detail in our companion narrative paper (den Besten and Känzig, 2026), our results are robust to reclassifying these borderline cases. In addition, our results are robust to excluding tariff episodes that coincided with potentially confounding non-tariff policy changes.

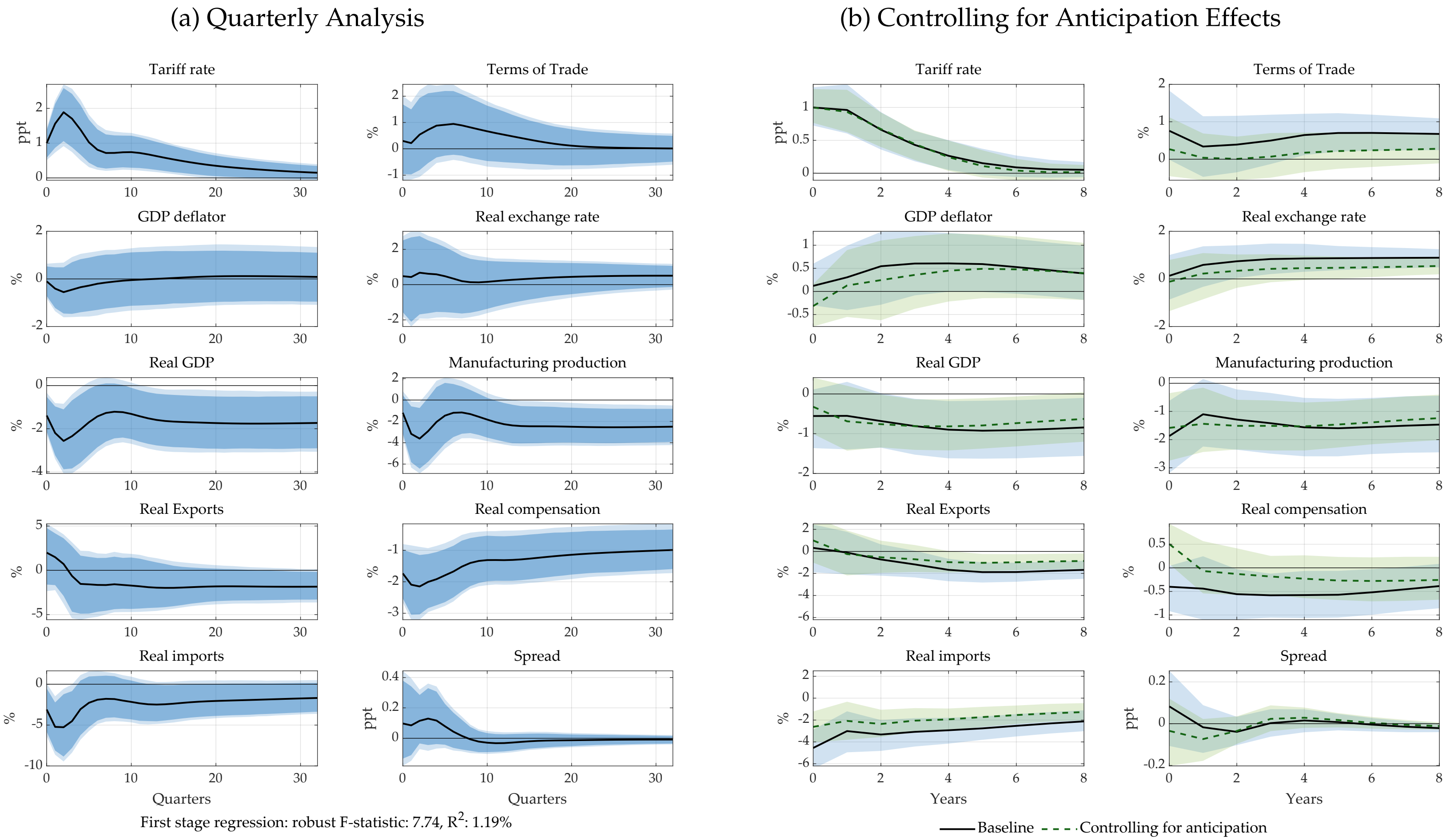
We conduct our baseline analysis at the annual frequency, due to data availability and reliability constraints over the long historical sample. At this frequency, tariff changes may coincide with other macroeconomic shocks within the same year. To address this concern, we construct a quarterly dataset extending back to the mid-nineteenth century. While the quarterly series are necessarily noisier given data limitations and the need for temporal disaggregation, the higher frequency substantially mitigates concerns about contemporaneous confounding shocks. Figure 11a reports the impulse responses estimated using the quarterly data. Reassuringly, the results closely mirror those obtained at the annual frequency. Imports, exports, output, and manufacturing activity all respond in a very similar manner. The main difference is a somewhat 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.

We conduct several additional robustness checks, reported in detail in the Appendix. First, we conduct a systematic leave-one-out (jackknife) exercise, re-estimating the model while sequentially excluding one tariff event at a time from the narrative series (see Appendix D.1). The resulting responses remain both qualitatively and quantitatively stable. Second, we implement a placebo exercise, assigning tariff shocks to dates without identified policy changes (see Appendix D.2). Reassuringly, randomly drawn placebo dates do not lead to systematic changes in trade, output, or prices. Overall, these exercises provide further support for the interpretation that our baseline results reflect the causal effects of exogenous tariff shocks rather than being driven by a small number of influential episodes or spurious confounders.

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

Figure 11: Accounting for Confounding Events and 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 external-instrument VAR model (2). Panel a: Quarterly VAR model, estimated on 1866Q1-2024Q4. Panel b: Excluding events with long signing-implementation gaps (over three months). 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.

gaps between signing and implementation dates. Figure 11b reports the resulting impulse responses. The responses are very similar to those obtained in the full sample. This finding suggests that anticipation effects are unlikely to be a major driver of 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 D. 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, reinforcing the conclusion that our results are not sensitive to specific implementation choices.

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.

We find that 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 importance of general-equilibrium effects: 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 more muted. Although higher tariffs raise trade barriers and input costs, inflation responds only modestly and imprecisely, reflecting the interaction of supply-side cost pressures with 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 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 and 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. From this perspective, the macroeconomic effects of recent tariff hikes are likely to hinge on the strength of these demand-side channels – particularly the extent of retaliation and the monetary policy response – rather than following mechanically from higher trade barriers alone.

References

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 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 S 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.

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

den Besten, Tamar and Diego R. Känzig (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 Pub. 0000. 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.

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.

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.

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

— (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.

Jordà, Òscar, Moritz Schularick, and Alan M. Taylor (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, E. (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.

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 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.

— (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. Chap. Series M 68-74.

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†
Diego R. Känzig‡

Contents

†Northwestern University. E-mail: tamardenbesten2029@u.northwestern.edu. Web: www.tamardenbesten.com.

‡Northwestern University, CEPR and NBER. E-mail: dkaenzig@northwestern.edu. Web: www.diegokaenzig.com.

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 a 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[3]
22	Smoot-Hawley Tariff Act	June 18, 1930	Increase	Distributional	Exogenous
23	Reciprocal Tariff Act	September 3, 1934[4]	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[5]	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[6]	Increase	Trade ideology; Political	Exogenous
35	2019 Trump Tariffs	May 10, 2019[7]	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

This appendix provides details on the construction of our annual and quarterly datasets.

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 1945–1889 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 1945–1889 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
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

Variable	Description	Source	Sample
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	Real general government revenue, deflated by GDP deflator	Müller et al. (2025)	1840–2024
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	1850–1960 1960–1913 1913–2024
Gold Stock	U.S. monetary gold stock, end of year, converted to fine ounces	Bordo (2006)	1879–1971

[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.

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
	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

Variable	Description	Source	Sample
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
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)	1955Q1–2025Q2
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)	1955Q1–2025Q2
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

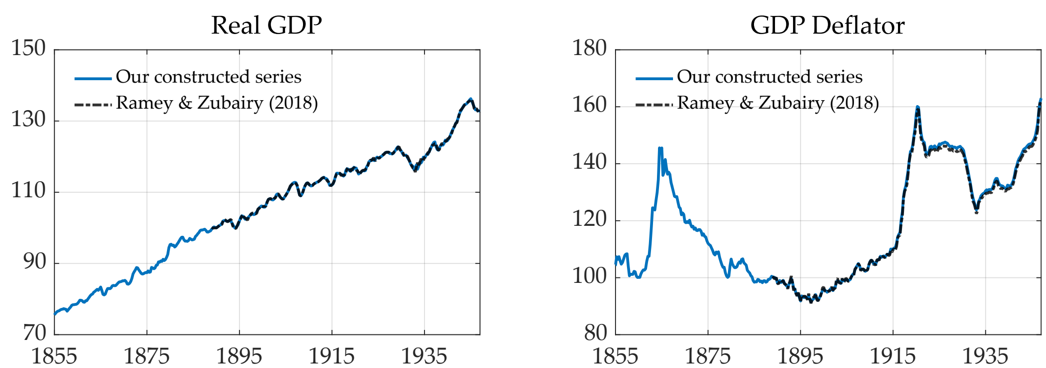
Variable	Description	Source	Sample
Manufacturing Production	Manufacturing production index		
	Annual manufacturing production index, interpolated using index of American business activity [M12003USM516NNBR][4]	NBER Macrohistory Database	1855Q1–1875Q1
	Annual manufacturing production index, interpolated 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
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 [5]	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		

Variable	Description	Source	Sample
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).

[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.

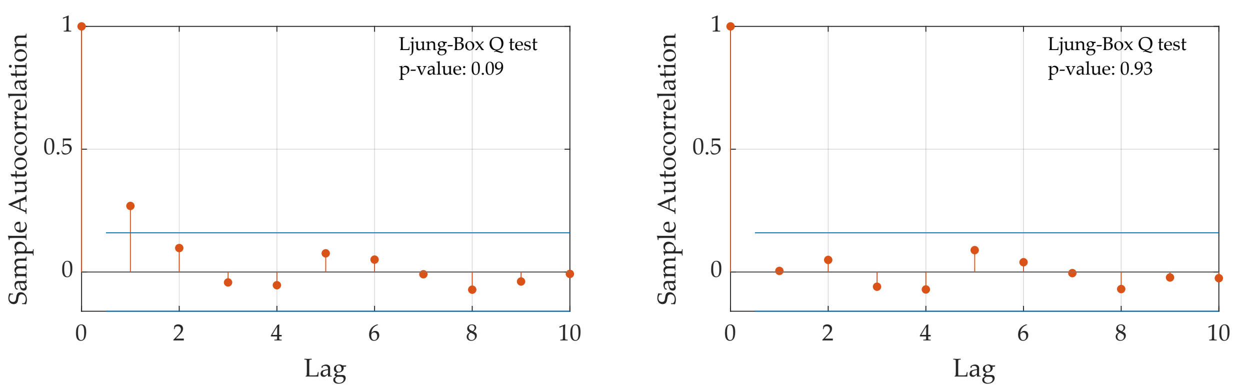
C. Instrument Diagnostics

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. The p-value on the instrument in Table 1 suggests that, conditional on the included macro variables, past values of the instrument do not significantly help predict its current value. This provides supportive evidence that the instrument behaves as a serially uncorrelated shock. Figure C.1a shows the autocorrelation function of the instrument. We can see that there is little evidence for autocorrelation. Only the first lag is borderline significant and the p-value for the Q-statistic testing the joint null that all autocorrelations up to 10 lags are zero is 0.24. Thus, we cannot reject the hypothesis that the instrument is serially uncorrelated.

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

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

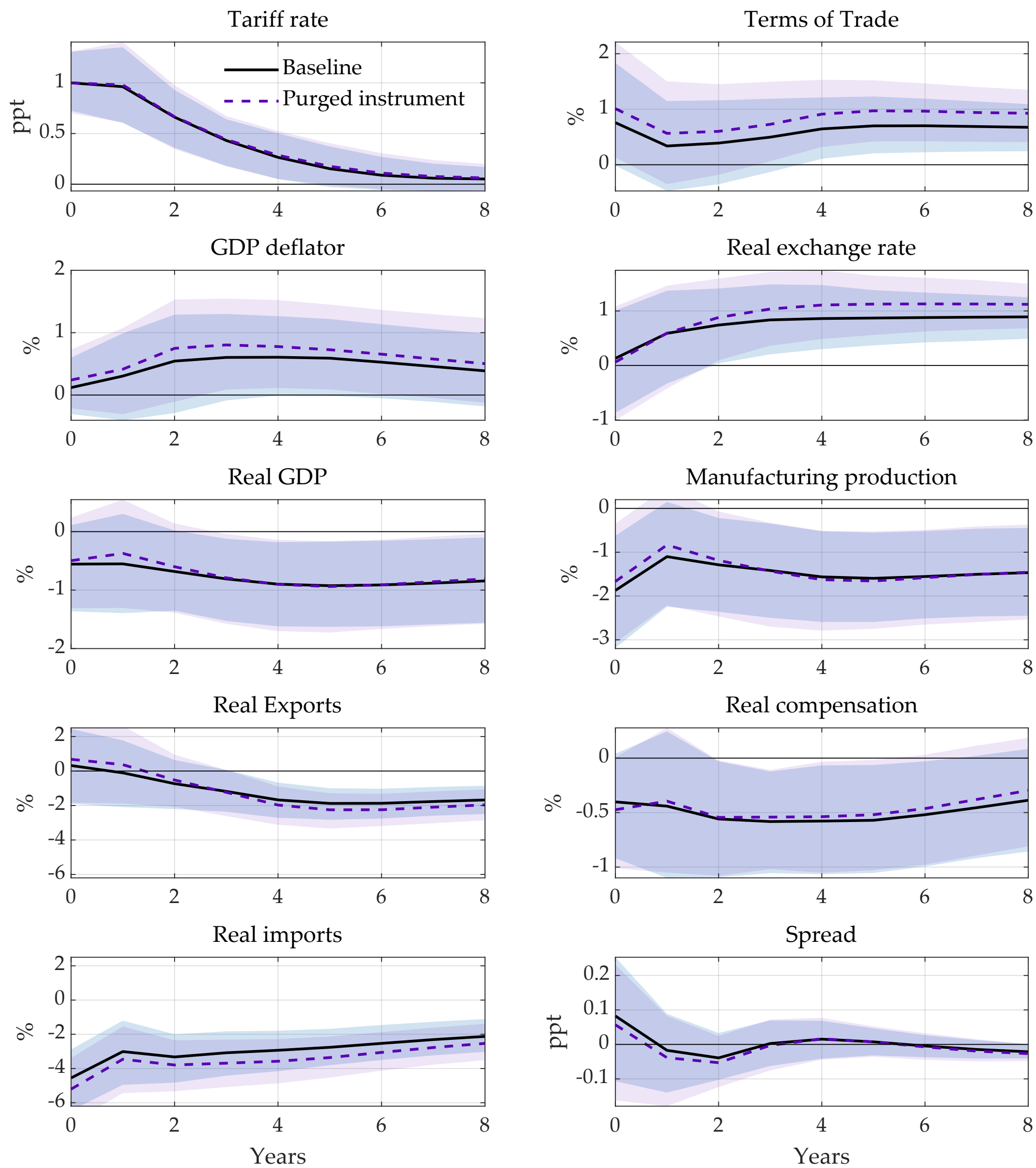


Nevertheless, we follow the approach in Miranda-Agrippino and Ricco (2021) and purge the distributed proxy, Z_t, of predictable dynamics as an additional check. We regress the instrument on its lag and use the residual, $Z_t^{\perp}$:

$$Z_t = \phi_0 + \phi_1 Z_{t-1} + Z_t^{\perp}$$

Figure C.1b shows that the autocorrelation of the purged instrument $Z_t^{\perp}$ is close to zero at all lags, confirming that the small amount of mechanically induced 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 (2). 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 reduces 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	0.69
Terms of trade	0.45
Real exports	0.67
Real imports	0.85
Macroeconomic variables	
Real GDP	0.68
Manufacturing production	0.53
Real compensation	0.97
GDP deflator	0.22
Producer price index	0.60
Financial variables	
Interest rate spread	0.86
REER	0.17
Real energy commodity price index	0.14
SP500	0.42
Government finances	
Real government debt	0.21
Real government revenue	0.87
Real government expenditure	0.25
Joint	0.77

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 influence the *timing* of tariff reforms indirectly if they help produce the unified governments 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	0.12	0.04
Tariff rate	0.94	0.10
Terms of trade	0.96	0.96
Real exports	0.89	0.55
Real imports	0.46	0.18
Macroeconomic variables		
Real GDP	0.75	0.82
Manufacturing production	0.95	0.11
Real compensation	0.82	0.49
GDP deflator	0.88	0.74
Producer price index	0.86	0.93
Financial variables		
Interest rate spread	0.88	0.41
REER	0.82	0.96
Real energy commodity price index	0.72	0.88
SP500	0.60	0.39
Government finances		
Real government debt	0.94	0.46
Real government revenue	0.94	0.63
Real government expenditure	0.96	0.11
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 the real energy commodity price shock in the same way, ordering the energy price series first in the VAR.

The fiscal policy shocks are taken from Romer and Romer (2010), who identify exogenous tax changes motivated by long-run policy objectives rather than current economic conditions. We use the military spending news shocks from Ramey and Zubairy (2018), which are 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
Fiscal 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. 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.

D.1. Instrument construction

Sensitivity to event selection. A core element of our identification strategy is the construction of a narrative instrument based on exogenous tariff events. In the previous section, we showed that the resulting shock series exhibits the desired properties for a valid instrument. We conduct an extensive sensitivity analysis in our companion narrative paper to assess the robustness of the baseline results to alternative event classifications (den Besten and Känzig, 2026). In particular, we examine whether the findings are sensitive to episodes in which fiscal considerations may have influenced tariff decisions, tariff reforms coincided with other major policy changes, or economic conditions deteriorated during legislative or negotiation processes, potentially affecting the timing or content of the final measures. We re-estimate the results under alternative treatments of these episodes and compare the resulting responses to the baseline specification. We summarize our findings below.

Excluding episodes in the early and mid-1870s, when budget surpluses and subsequent deficits may have created pressures for tariff reforms, leaves the estimated responses largely unchanged.

In a few episodes, tariff legislation was enacted alongside other major policy measures, raising the possibility that estimated tariff effects may partly capture the influence of contemporaneous legislation. Excluding tariff revisions that coincided with broader fiscal reforms, such as the introduction of the federal income tax, or with multilateral trade agreements which also included reductions in non-tariff barriers, yields impulse responses that remain closely aligned with the baseline.

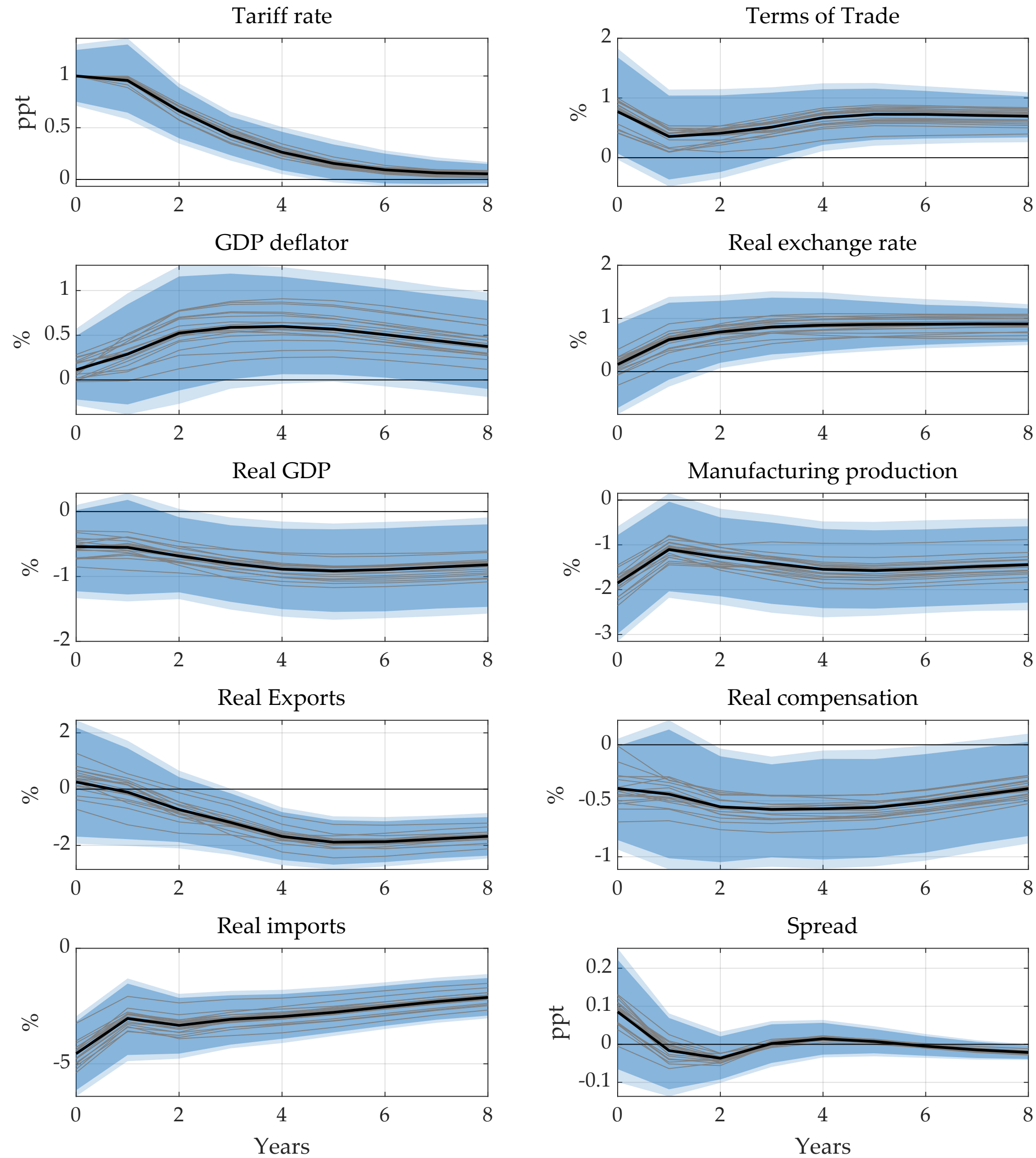
Finally, we consider episodes in which economic conditions deteriorated during legislative or negotiation processes, potentially influencing the timing or content of tariff reforms. For example, the Smoot–Hawley Tariff was initially motivated by distributional concerns, particularly extending protection to agriculture, but economic conditions worsened during the legislative process after the bill's initial passage. Excluding this episode leaves the estimated responses largely unchanged.

Taken together, the sensitivity analysis suggests that our baseline results are not driven by the treatment of potentially contentious episodes.

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 D.1 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.

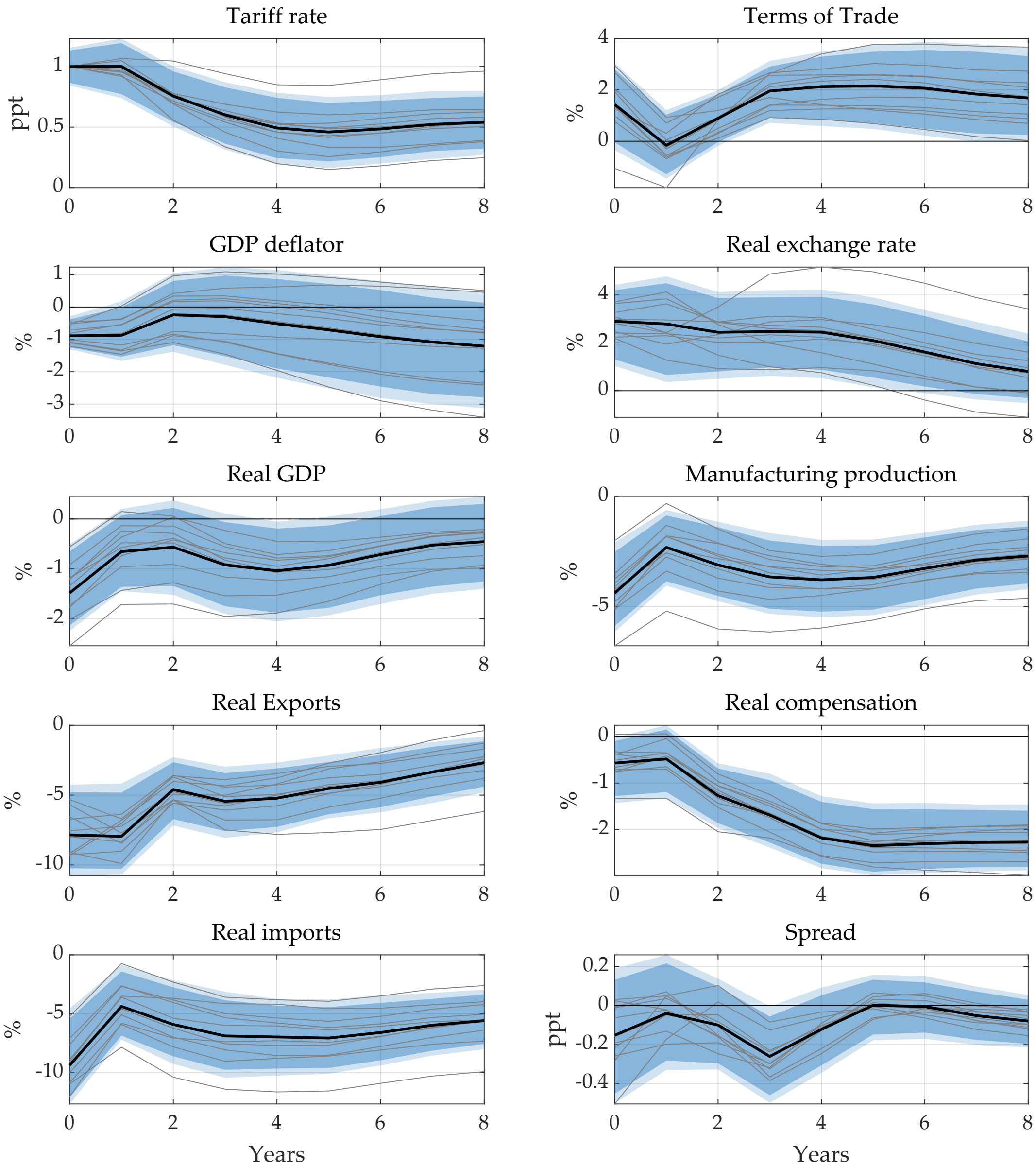
We repeat the jackknife exercise for the postwar sample and plot the results in Figure D.2. In contrast to the full historical sample, the responses are noticeably less stable across jackknife iterations, reflecting the limited number of major tariff events and the resulting sensitivity to the exclusion 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 D.1: 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 (2), 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.

Figure D.2: Results from Jackknife Exercise for the Postwar Sample

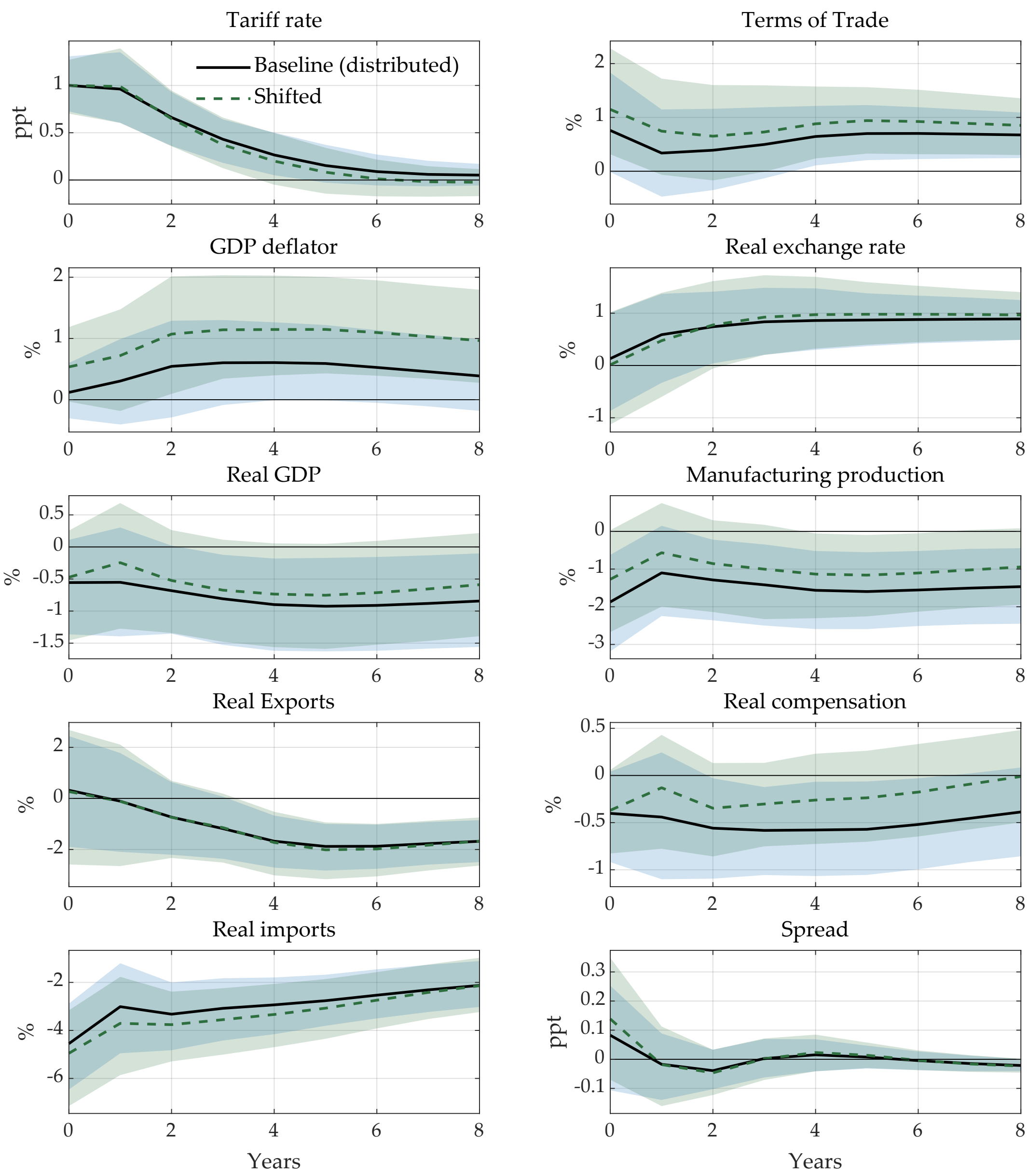


Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated based on the external-instrument VAR model (2), excluding one event in the instrument at a time. Sample: 1946-2024. Lag order: 2. Solid line and shaded areas: baseline 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.

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 D.3: 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 (2). 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 D.3 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.

Size of the shocks. The key motivation behind our qualitative narrative shock series is that while the timing of a tariff event is plausibly exogenous, the size of the tariff change may be endogenous. Furthermore, accurately measuring the size of legislative tariff changes is 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. Thus, we use narrative sign dummies as our baseline instrument, capturing the exogenous timing and direction of tariff changes.

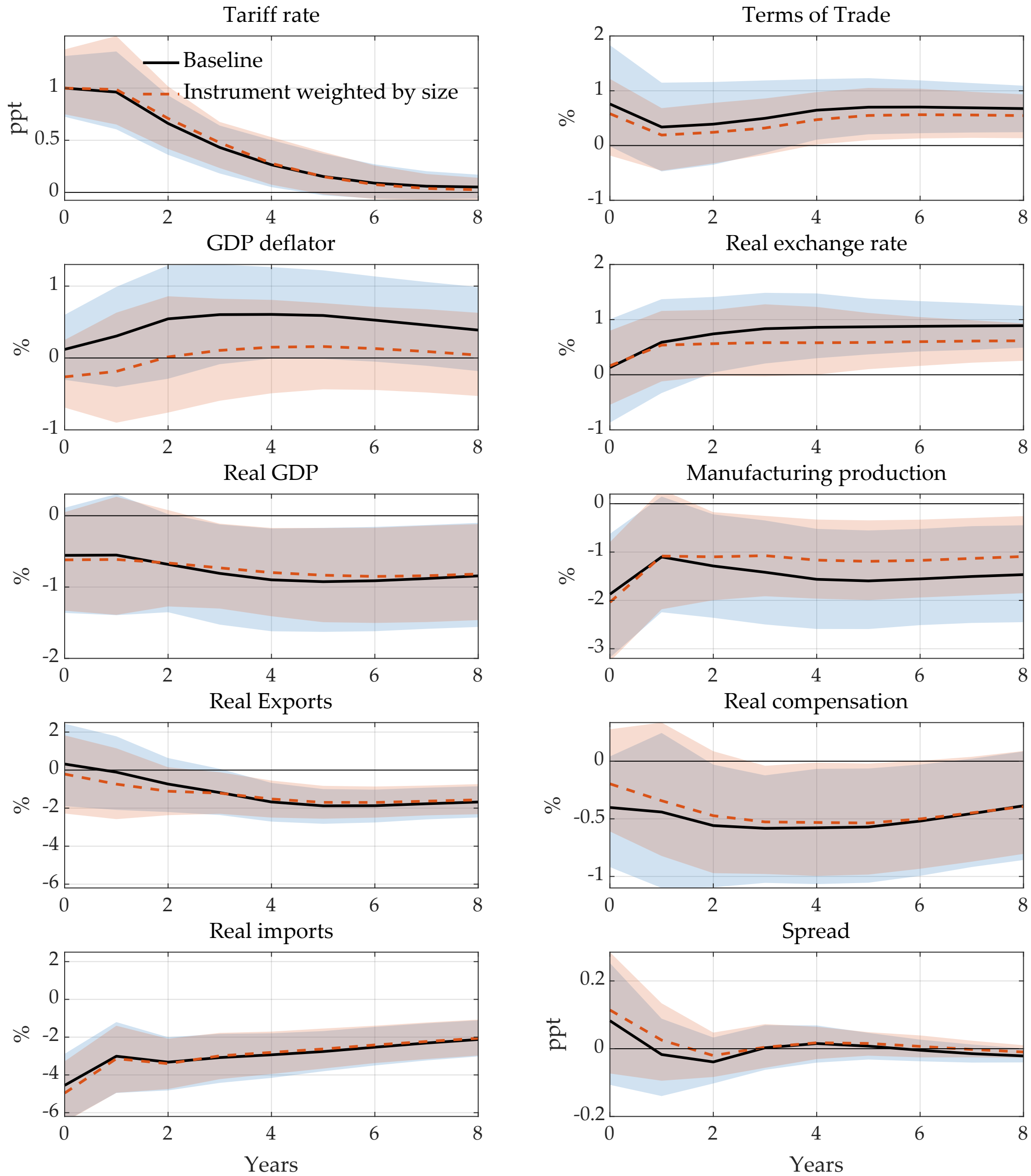
However, information on the sign of an event is relatively weak, which may lead to weak instrument problems. Even though the first stage F-statistic does not point to a weak instrument problem at hand, we explore how much more power we can get when including quantitative information as well.

How to best measure the size of a given tariff shock? This is a challenging task because shifts in imports and especially import prices affect the average tariff rate contemporaneously, making it an imperfect measure of the size of tariff changes. Ideally, we would hold the composition of imports and their prices fixed, and use product-level tariff rates to calculate the policy-driven change in tariff revenue for each reform. While such a measure could be constructed 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 D.4 shows the results. Overall, the responses turn out to be similar. The main 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-

Figure D.4: 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 (2) 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 change. Light shaded areas: 90 percent confidence bands based on moving-block bootstrap.

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 5.

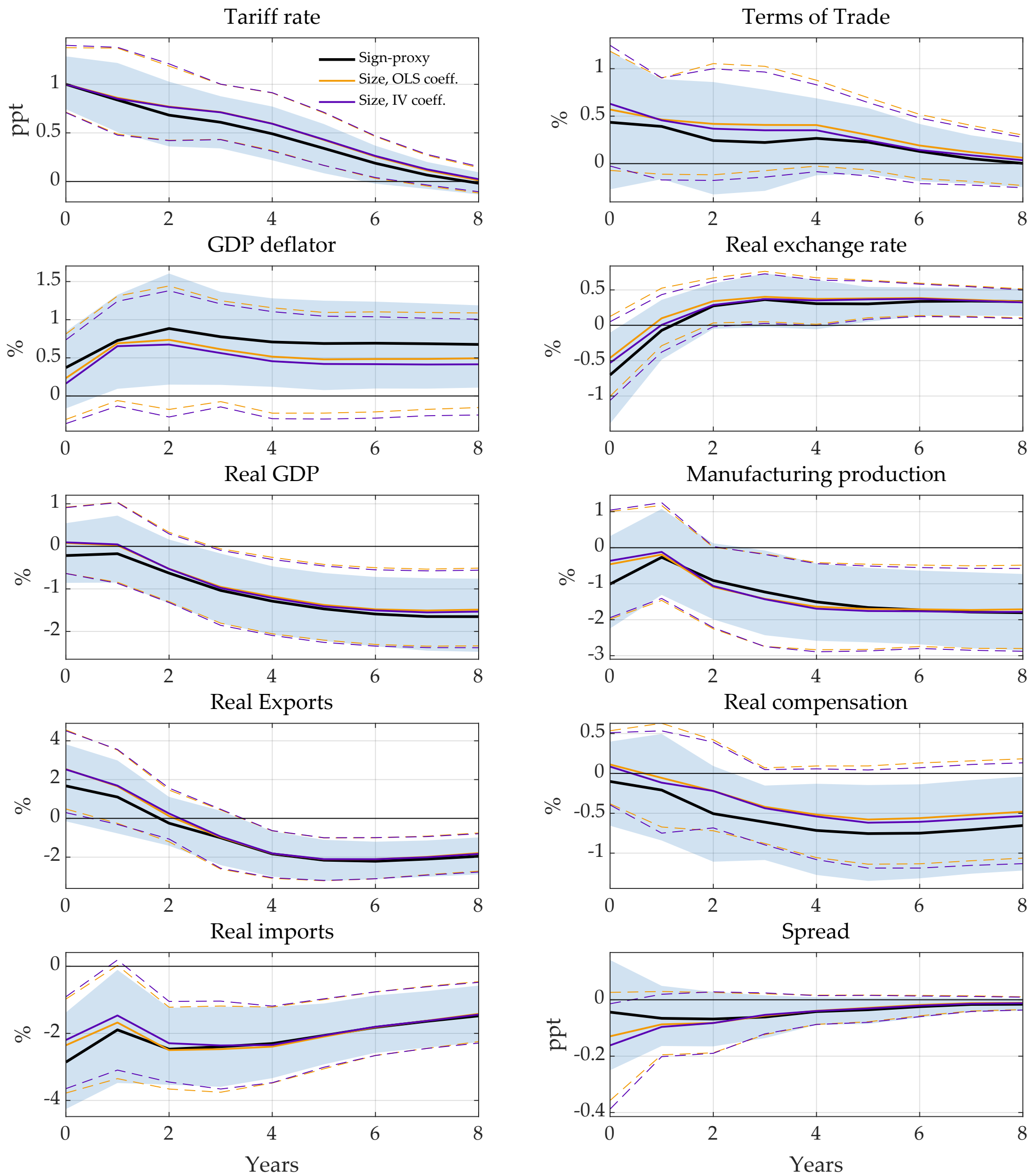
As an alternative size measure, we use estimates by Irwin (1998). To separate the effects of import prices from the effects of legislative tariff changes on the average tariff rate, Irwin (1998) regresses the average tariff rate on dutiable imports on the log of the import price index and dummy variables 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 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 D.5 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 and the exclusion of later tariff changes. Importantly, the results using the size-weighted instrument turn out to be very similar as well, even though the deflator response is again somewhat weaker. This is true when we use the OLS or the IV coefficients in Irwin (1998) (Columns 2 and 3 of Table 2, respectively). The F-statistics for the size-weighted instrument turn out to be substantially higher: close to 30 when using the OLS coefficients and close to 40 when using the IV coefficients. Despite this increase in first-stage strength, the confidence bands are comparable to those obtained with the sign-based instrument, suggesting that the narrative sign instrument is sufficiently strong to support valid inference.

Taken together, these results provide further support for using the narrative signed shocks as the baseline instrument, as their identifying assumptions are weaker and more plausibly satisfied while delivering comparable first-stage strength to that of size-weighted instruments.

Figure D.5: Incorporating the Size of Tariff Changes Based on Irwin (1998)



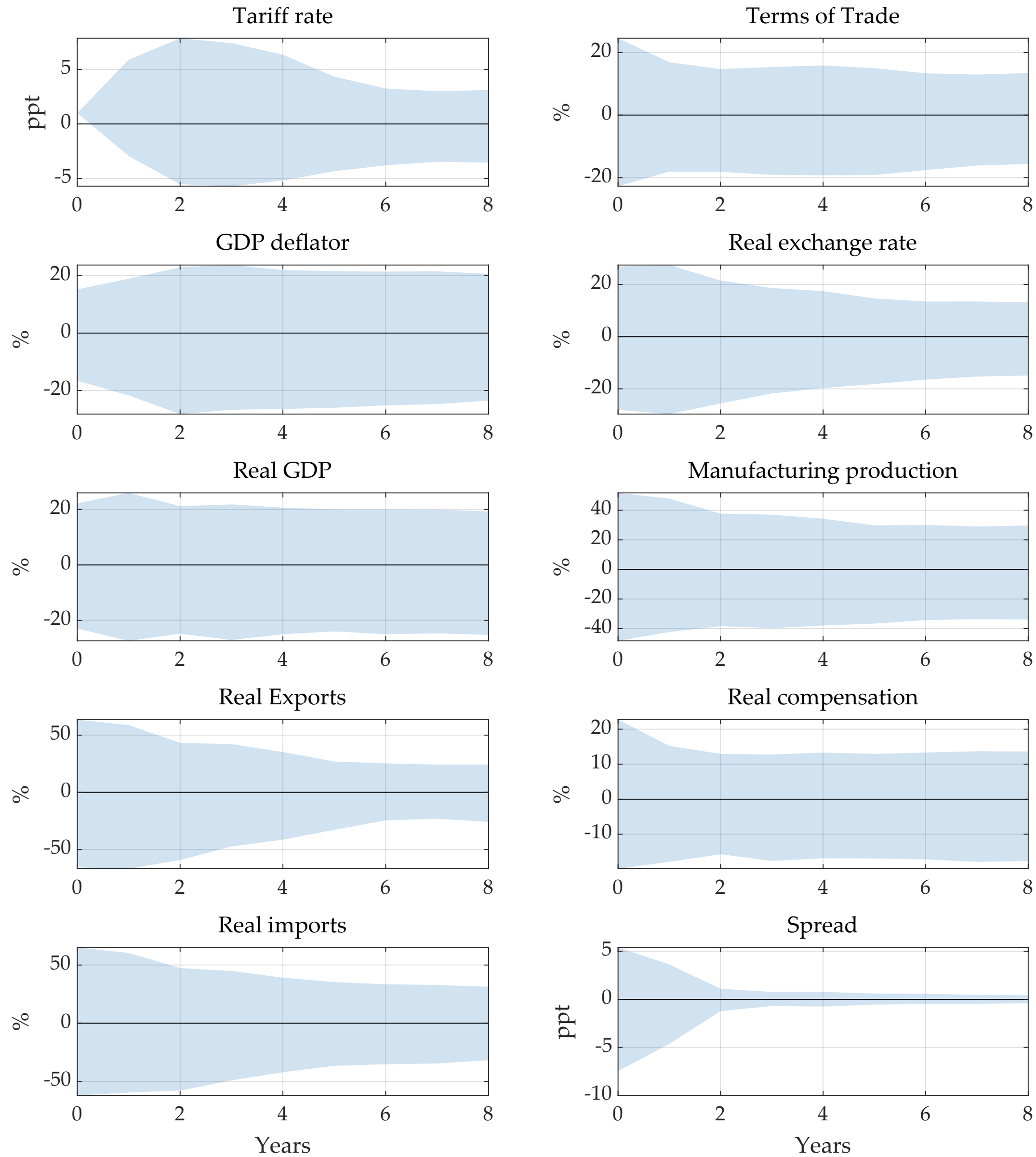
Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated based on the external-instrument VAR model (2) using sign-based and size-weighted narrative tariff shocks. Sample: 1865–1967. Lag order: 2. Solid black line: baseline 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.

D.2. Placebo exercise

In the next exercise, we show that randomly-drawn placebo dates do not generate systematic macroeconomic responses. We assign random non-event dates to each of the 20 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 D.6 shows the 5th and 95th percentiles of the placebo responses. Assigning random non-event dates to the 20 exogenous tariff events yields impulse responses that do not resemble the systematic macroeconomic effects observed for the actual tariff shocks. The wide bands are consistent with the volatility expected when the external instrument provides little explanatory power for the tariff shock. Importantly, they do not exhibit any systematic dynamics. Overall, this evidence suggests that the external instrument VAR approach is not picking up spurious correlations in the data.

Figure D.6: Placebo Responses

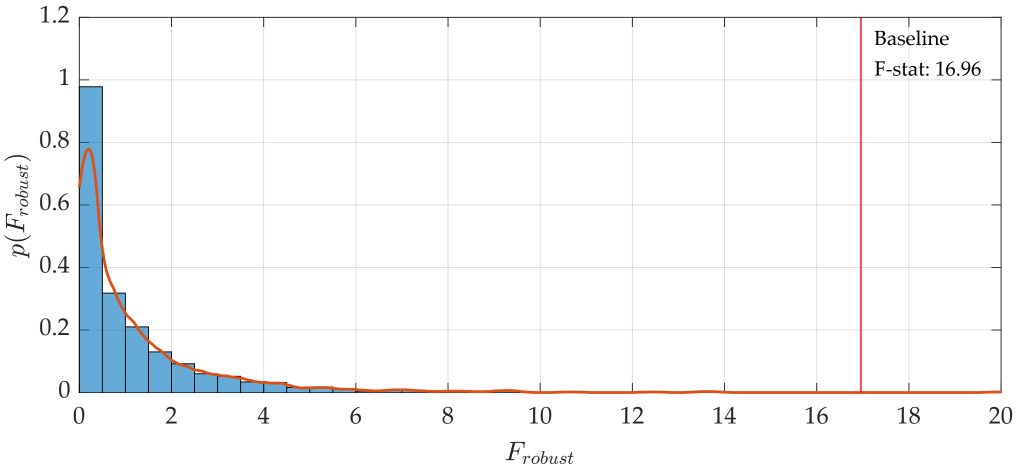


Notes: Shaded areas represent 5th and 95th percentiles of the placebo responses.

Another way to assess this is 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 D.7 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.

Figure D.7 displays the histogram of robust first-stage F-statistics. 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. These results confirm that the placebo series are weak instruments, consistent with the highly variable impulse responses in Figure D.6.

Figure D.7: F-statistics of Placebos



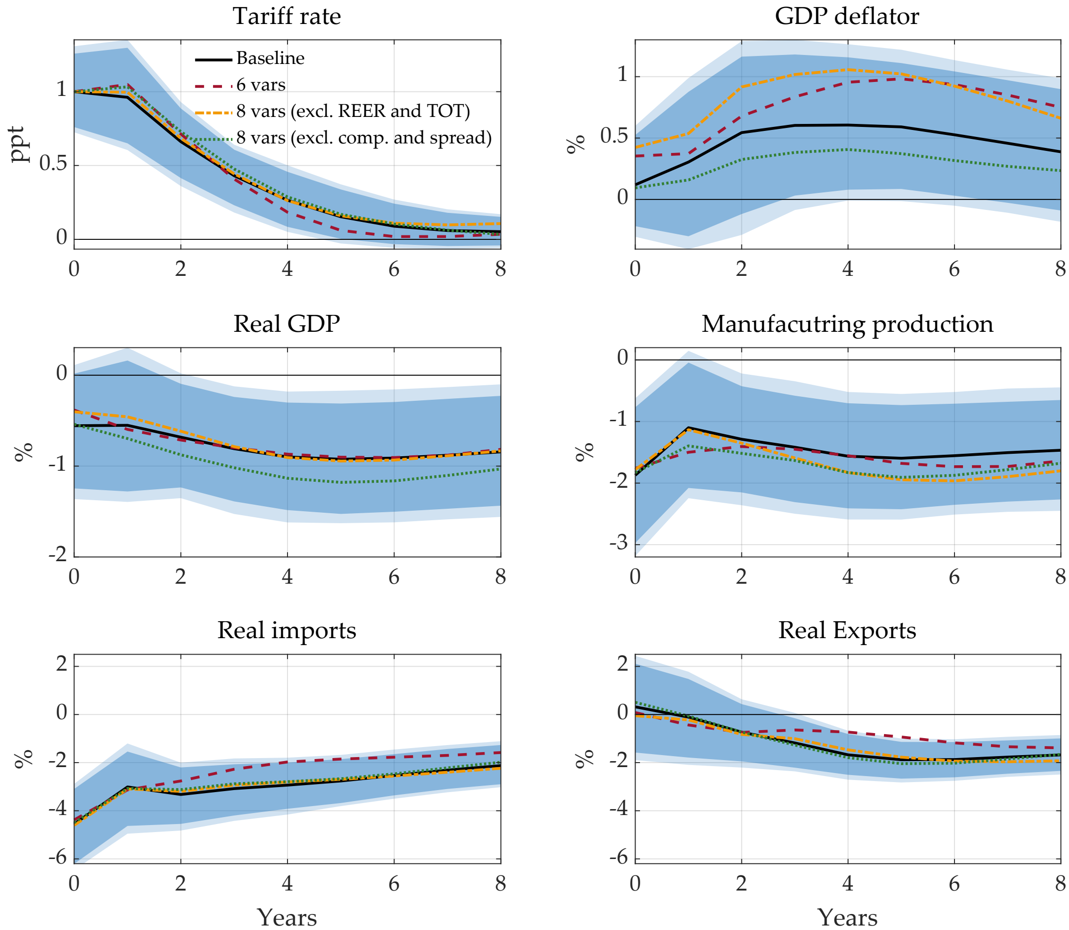
Notes: The distribution is based on the robust first-stage F-statistics obtained from 1,000 placebo instruments. The red vertical line indicates the value of our baseline first-stage F-statistic.

D.3. Specification choices

In the following section, we examine the sensitivity of the results with respect to specification choices. The baseline VAR includes 10 variables, which is relatively large given the annual 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 D.8 shows the point estimates of this 6-variable VAR, which remain consistent with those from the baseline specification.

To further examine sensitivity to variable selection when reducing the VAR dimension, we also estimate an 8-variable model. In particular, 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 D.8. The results remain robust to both the size of the VAR and the specific variables removed.

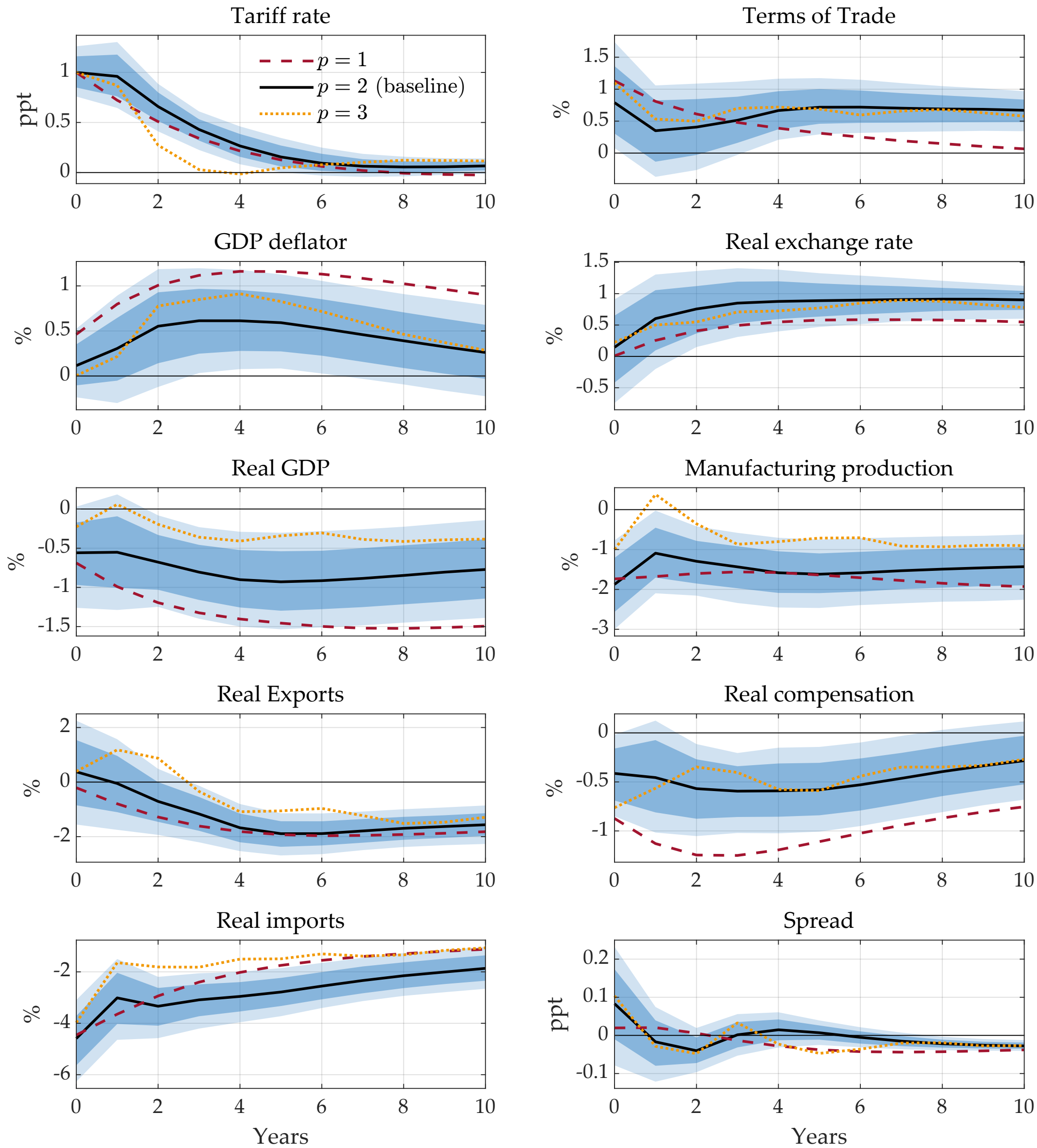
Figure D.8: Responses from Smaller VAR



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 with respect to alternative lag structures. Figure D.9 plots the baseline specification with two lags together with the point estimates obtained when using one or three lags. The magnitude and timing of the responses are somewhat sensitive to varying the lag order, with real GDP in particular exhibiting a weaker decline when additional lags are included. However, the qualitative direction remains consistent with our baseline.

Figure D.9: Sensitivity to Lag Orders



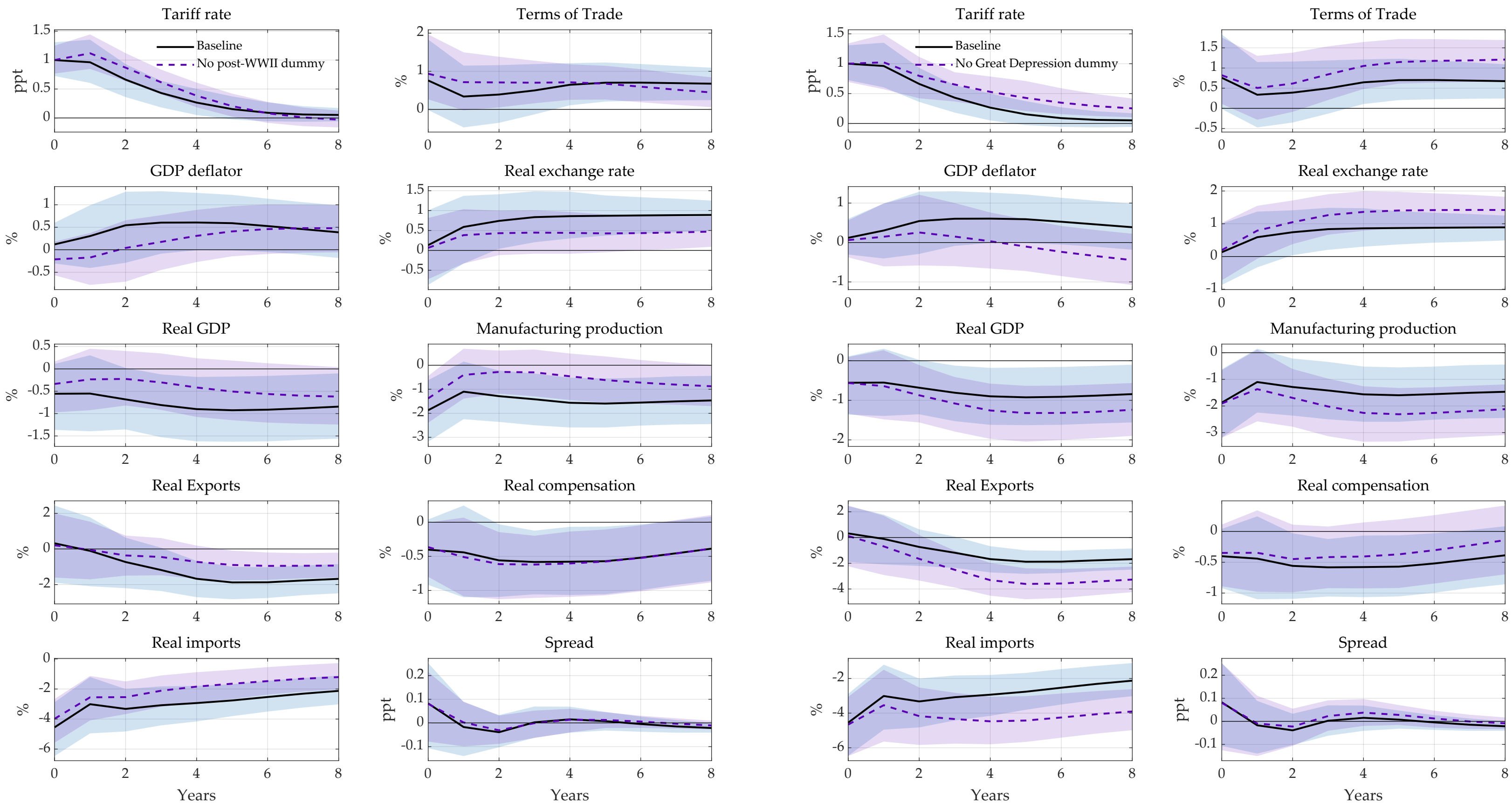
Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated based on the external-instrument VAR model (2) with varying lag orders. Solid black line and shaded areas: baseline responses with 90 and 95 percent confidence bands based on moving-block bootstrap. Dashed and dotted 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 D.10: Sensitivity with Respect to Deterministics I

(a) Excluding the Post–World War II Dummy

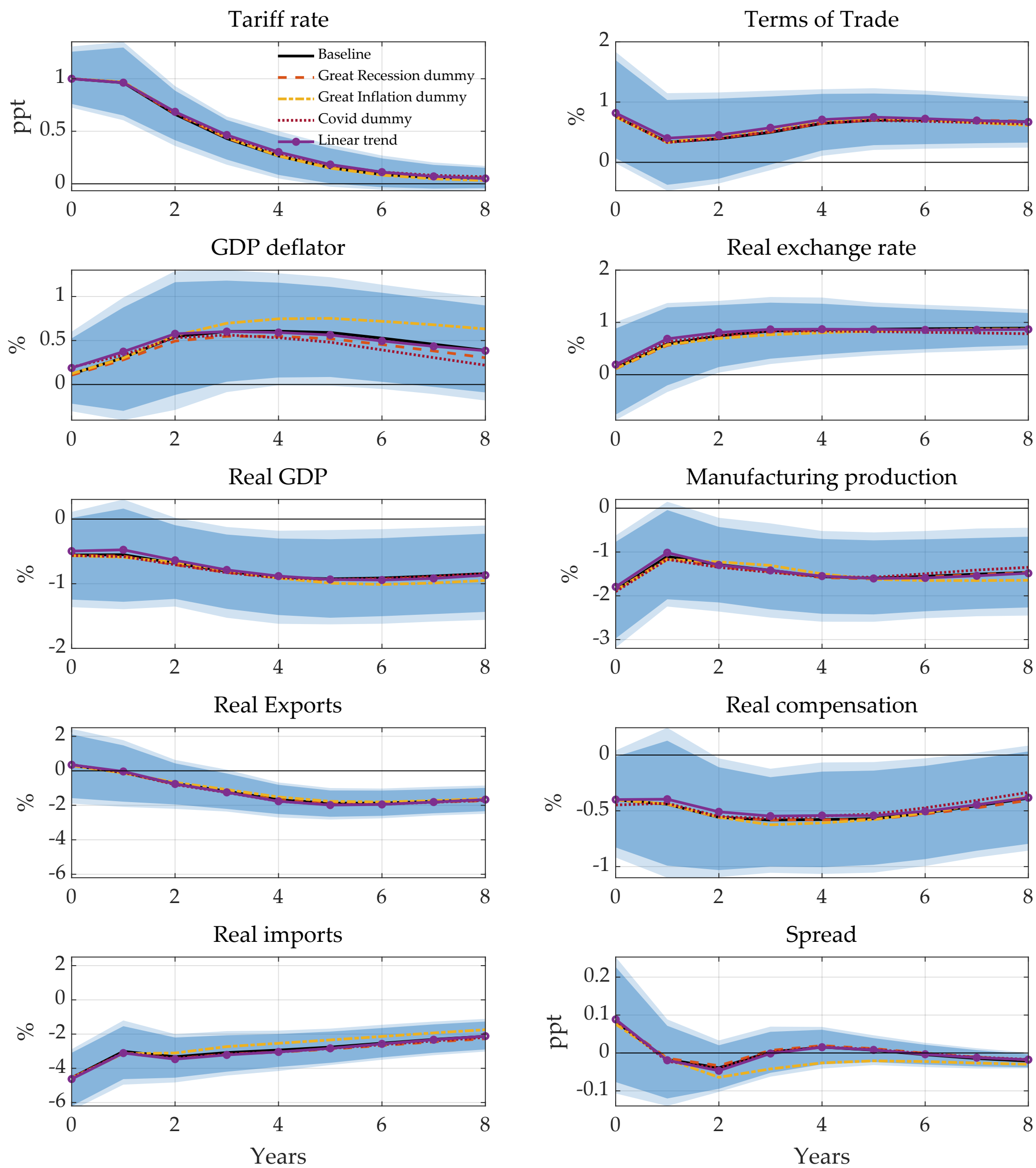
(b) Excluding the Great Depression Dummy



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 (2) 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 D.10, 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 D.11: Sensitivity to Deterministics II

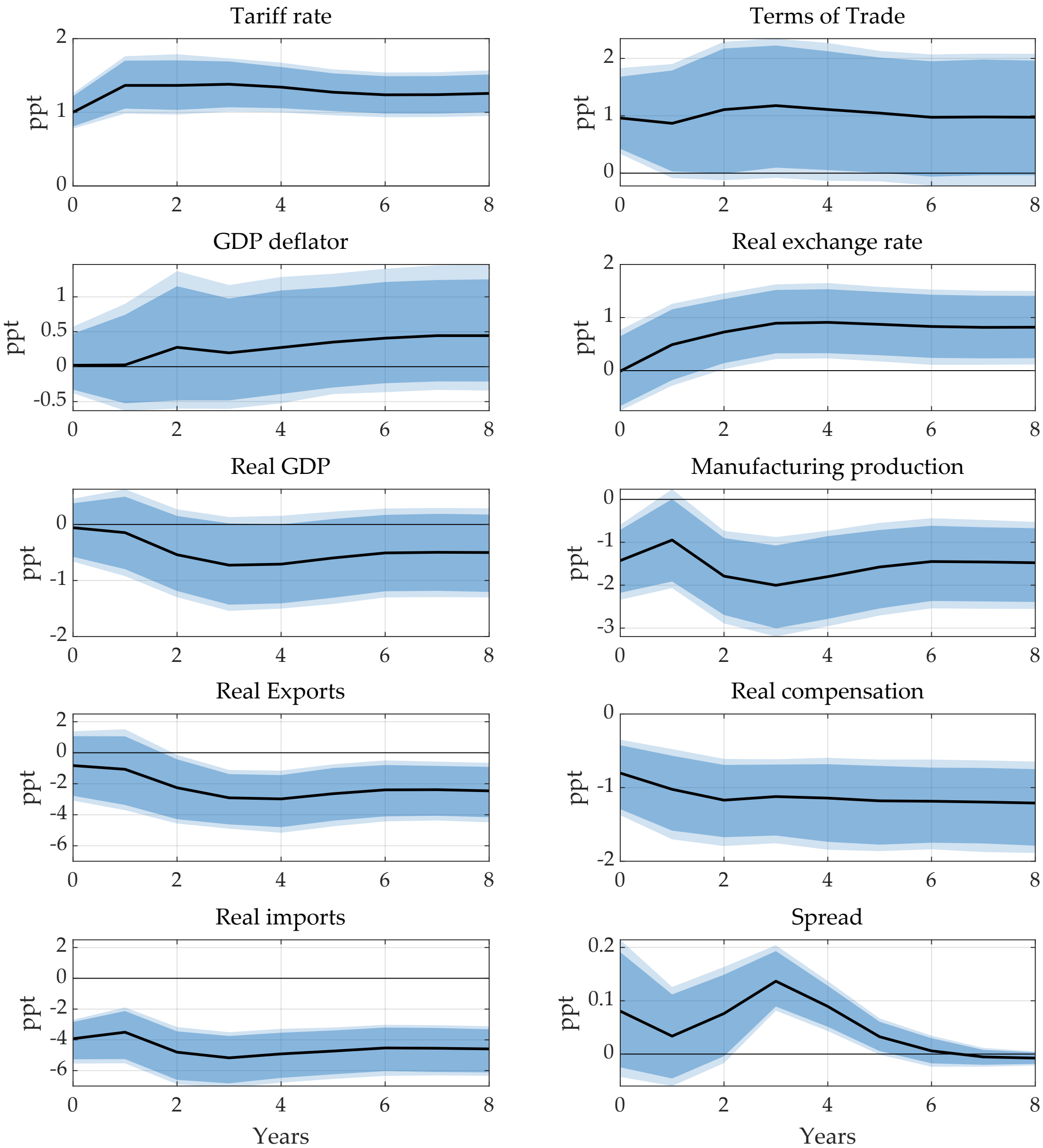


Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated based on the VAR model (2) when controlling for additional deterministic variables. Solid black line and shaded areas: baseline responses and 90 and 95 percent confidence bands based on moving-block bootstrap. Dashed and dotted lines: responses when including additional deterministic controls.

Figure D.11 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.

Finally, we estimate the VAR model in differences instead of log-levels. Figure D.12 shows that the responses remain qualitatively similar.

Figure D.12: Specification in Differences



First stage regression: robust F-statistic: 28.15, R^2: 8.59%

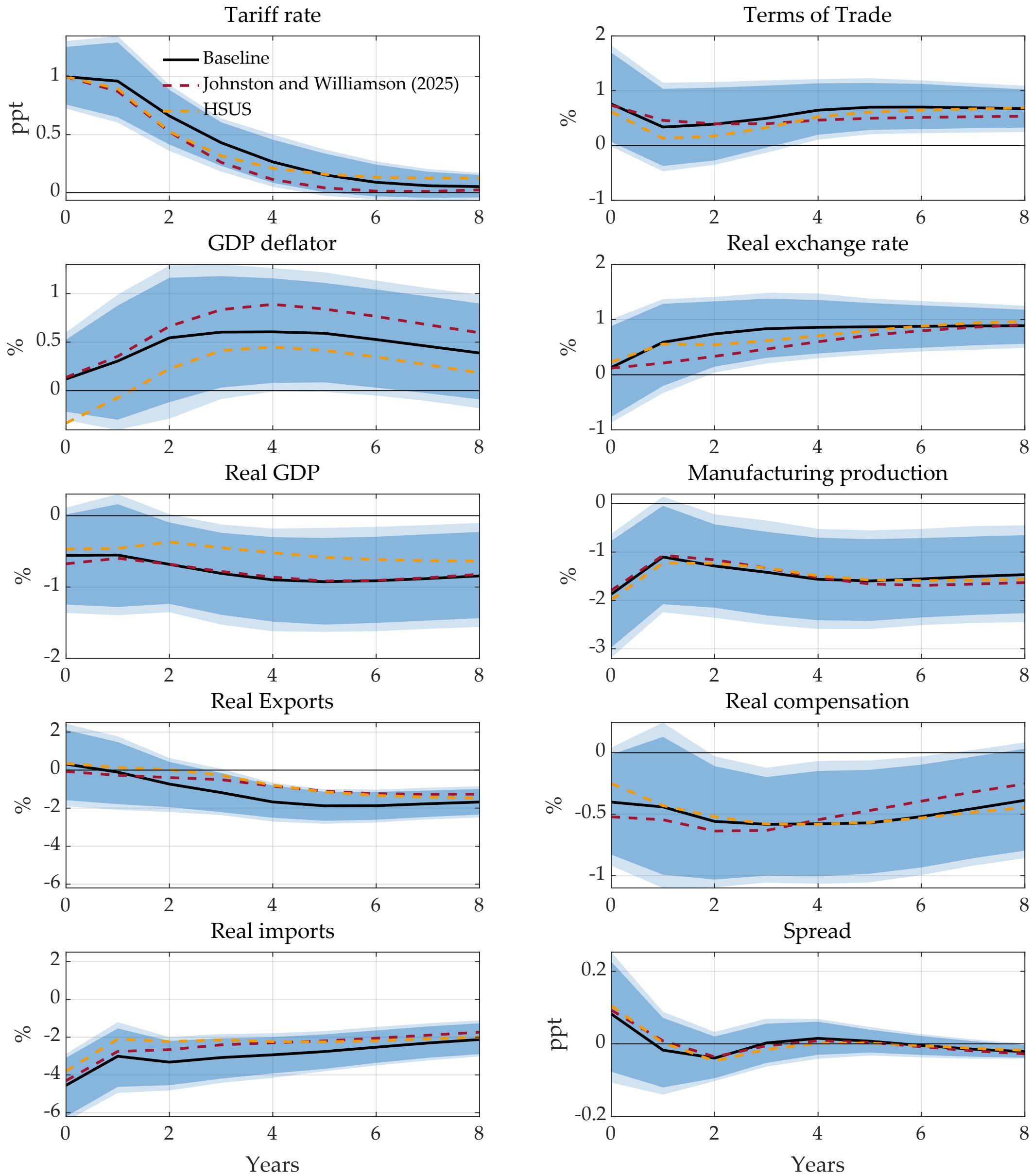
Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, based on a VAR estimated in differences with the narrative tariff shock series used as an instrument. Lag order: 2. Solid line: point estimate. Dark and light shaded areas: 90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

Comparison to Barnichon and Singh (2025). Our analysis differs from Barnichon and Singh (2025) along several important dimensions. First, our narrative instrument is broader, covering a larger set of tariff events (21 compared to 9), and we differ in the timing of these events by distributing tariff changes over the relevant years, rather than assigning them to a single date. Second, our empirical specification includes deterministic controls designed to account for major structural breaks and large confounding episodes – most notably the World Wars and the Great Depression – which play a central role in long-run historical data. Third, we control for a richer set of macroeconomic variables, allowing us to better isolate the effects of tariff shocks from other contemporaneous influences. Taken together, these differences contribute to the more nuanced price responses we document, in contrast to the sharper deflationary effects reported in Barnichon and Singh (2025).

D.4. Alternative variables and sources

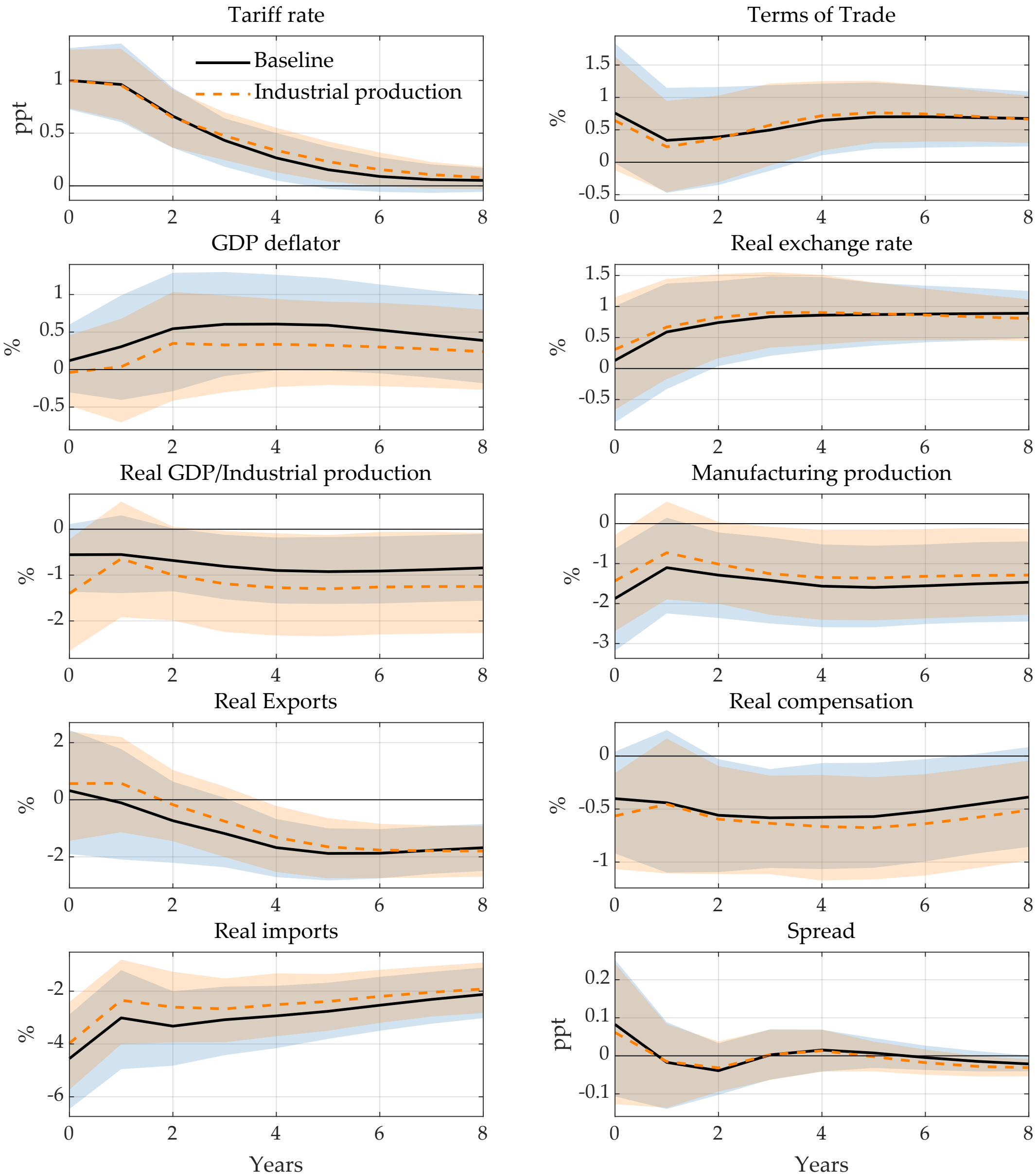
Historical data is subject to measurement challenges. In this appendix, we show that our results are robust to alternative data and measurement choices. Specifically, we re-estimate the VAR using different series for GDP and the GDP deflator (Figure D.13), replace GDP with industrial production (Figure D.14), and deflate exports and imports using their respective price indices rather than the GDP deflator (Figure D.15). 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 D.13: Responses using Alternative Measures of Output and Prices



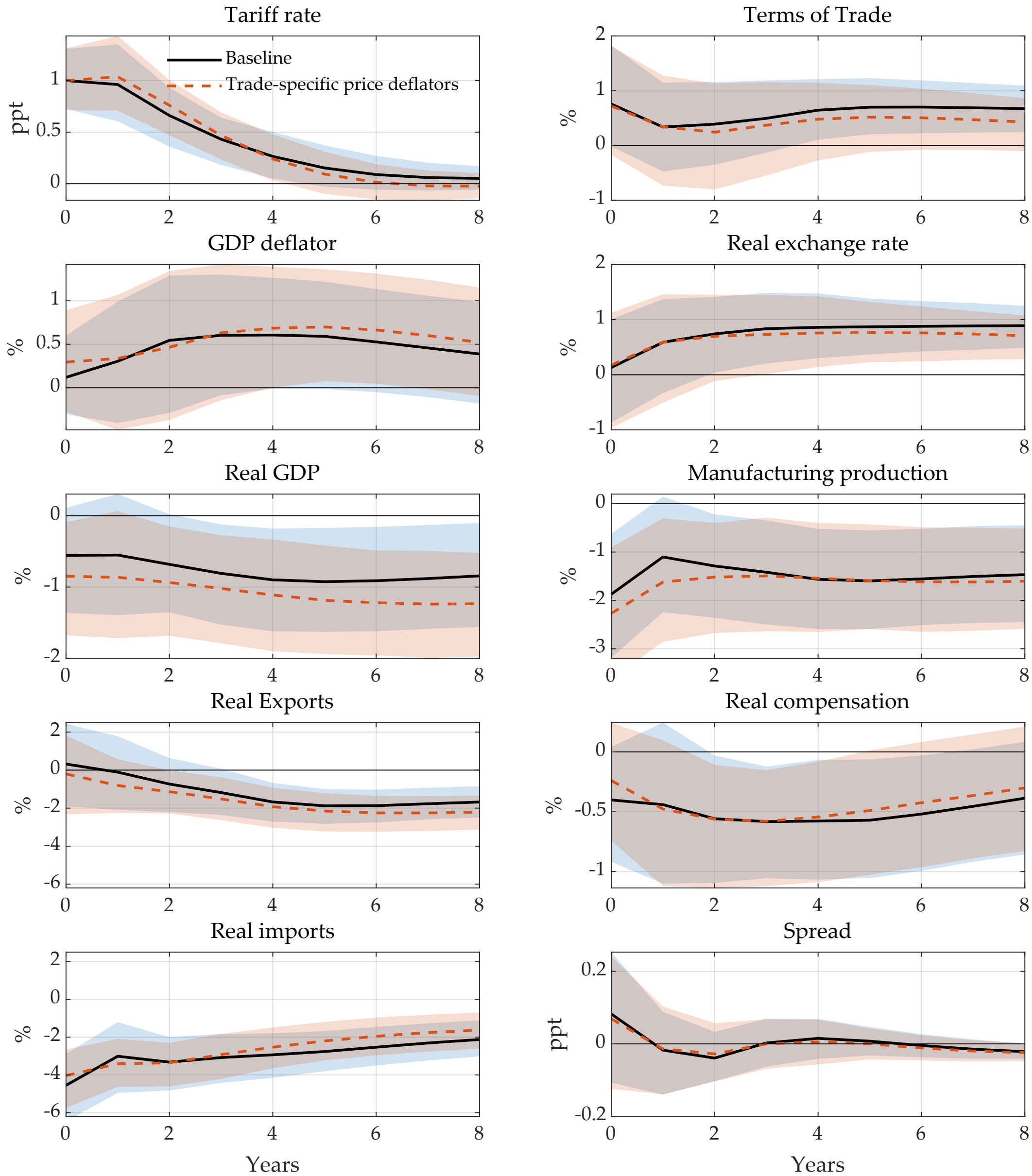
Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated based on the external-instrument VAR model (2) using alternative GDP measures. Lag order: 2. Solid black line and dashed areas: baseline responses with 90 and 95 percent 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 D.14: 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 (2) 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 D.15: Responses Using Trade-Specific Price Deflators



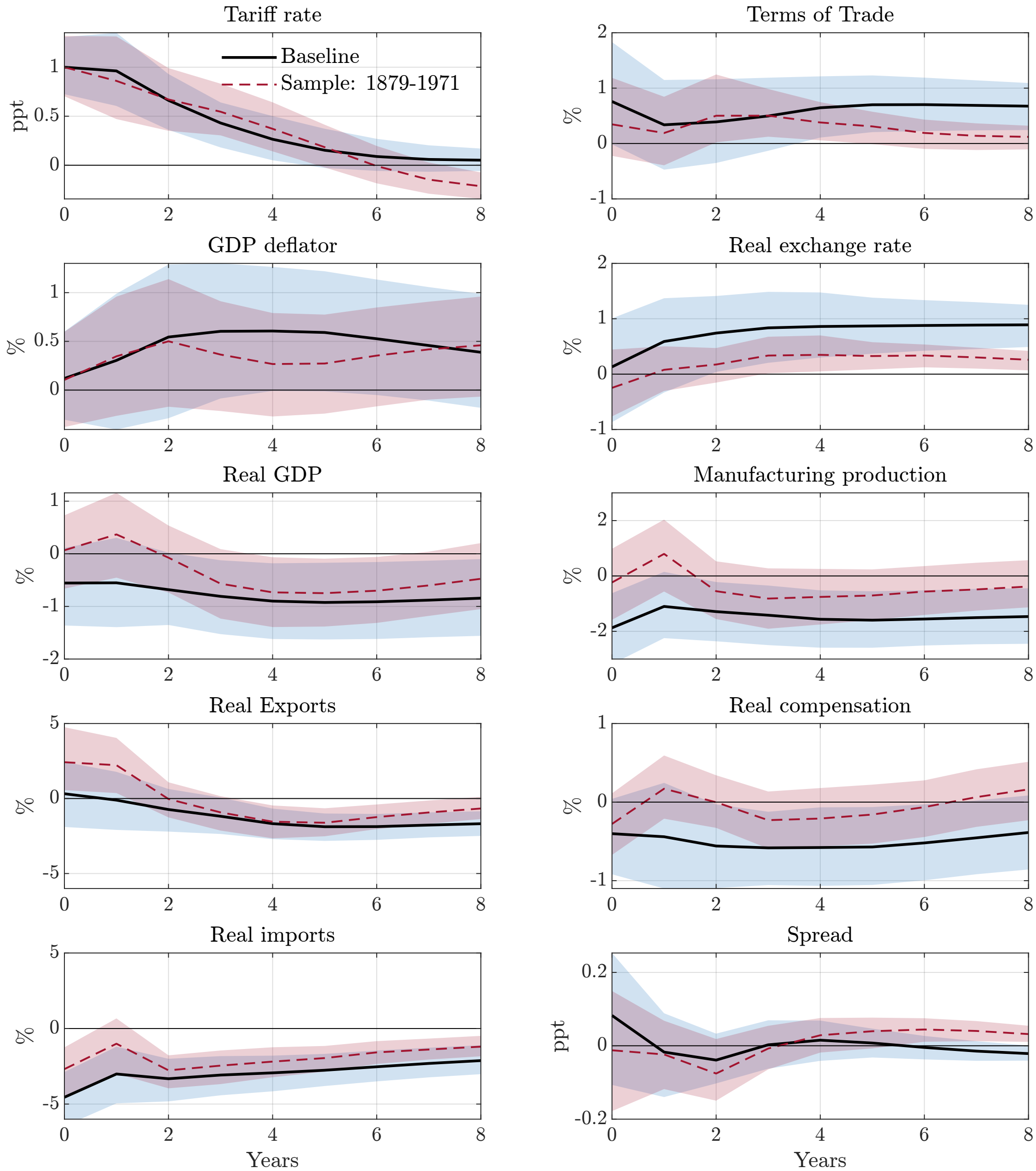
Notes: Impulse responses to a tariff shock, normalized to increase the dutiable tariff rate by 1 percentage point on impact, estimated based on the external-instrument VAR model (2) using different deflators for the trade variables. Lag order: 2. Solid black line: baseline responses. Dashed red line: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

E. Additional Results

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

Era of gold-backed monetary systems. To shed light on the role of monetary policy, we estimate our model on the sample featuring gold-backed monetary systems (1879–1971). Figure E.1 compares the responses over this sample to our baseline responses.

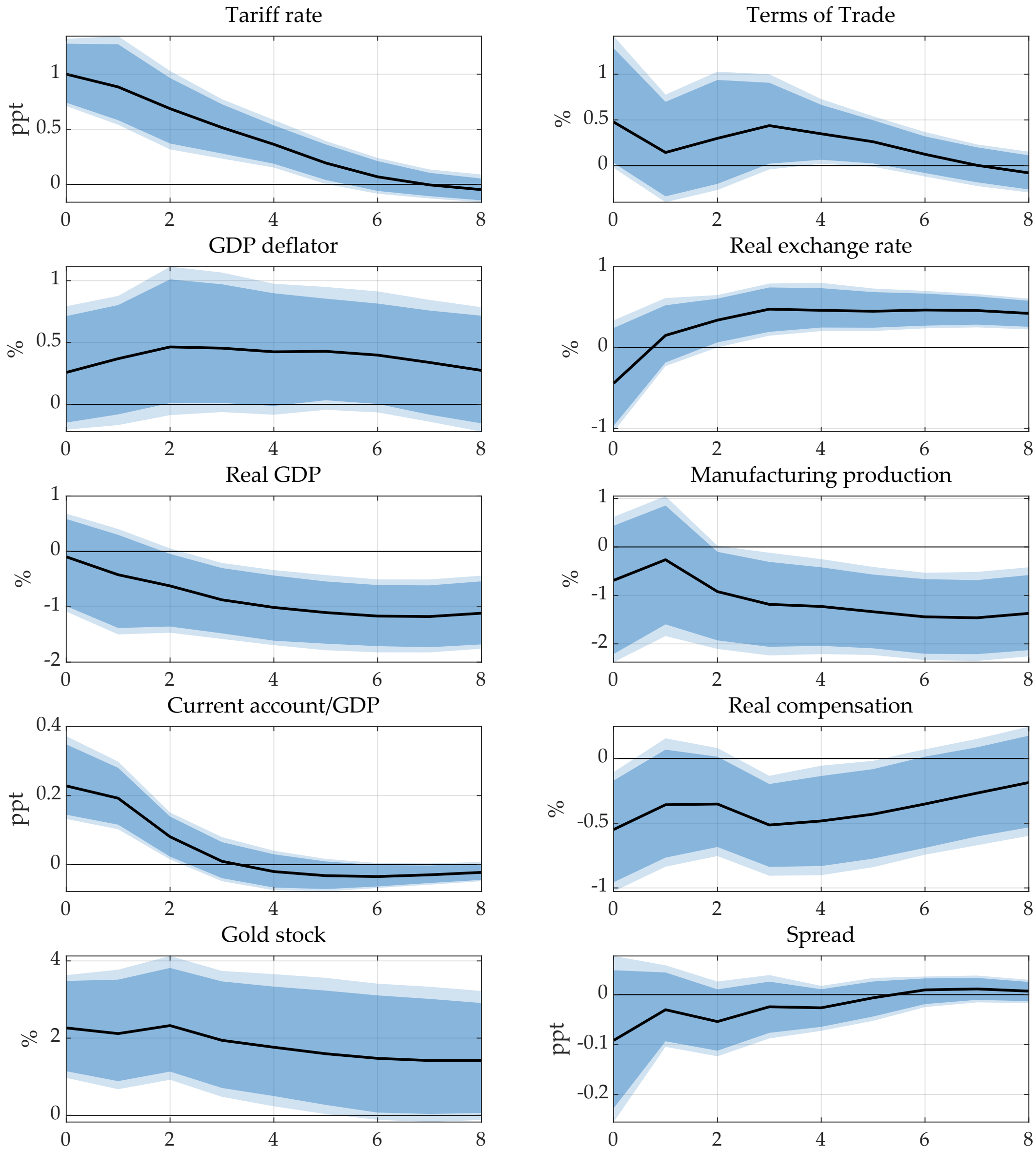
Figure E.1: Responses for Era of Gold-Backed Monetary Systems vs. Full Sample



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 (2) 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 E.2.

Figure E.2: 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 (2), 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.

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.

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.

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.

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.

— (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*.

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.

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.

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. Chap. Series M 68-74.

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.