

# Measuring Shortages Since 1900\*

Dario Caldara      Matteo Iacoviello      David Yu

October 9, 2024

## PRELIMINARY DRAFT

### Abstract

This paper introduces a monthly shortage index for the United States from 1900 to 2023, constructed using a sample of 25 million newspaper articles. The index measures the intensity of shortages of labor, materials, goods, and energy by calculating the proportion of articles discussing shortages each month. The resulting index reveals significant variation in shortage intensity over time, with notable peaks during periods of economic turmoil and wars. We explore the relationship between the shortage index and key economic indicators, discussing potential applications for researchers, policymakers, and businesses. Increases in shortages that are orthogonal to standard demand, supply and commodity price shocks generate persistent inflationary effects.

**KEYWORDS:** Shortages; Textual Analysis; Inflation.

**JEL CLASSIFICATION:** C43, E32, N11, N12.

---

\*Corresponding author: Matteo Iacoviello ([matteo.iacoviello@frb.gov](mailto:matteo.iacoviello@frb.gov)). The data on shortages presented in this paper are available at the <https://www.matteoiacoviello.com/shortages.html>. Caldara and Iacoviello are at the Federal Reserve Board. Yu is at UCLA. The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of anyone else associated with the Federal Reserve System.

# 1 Introduction

Shortages, defined as a lack of sufficient supply to meet demand, have been a recurring feature of economic life throughout the 20th and early 21st centuries. The hallmark of economics is the study of how to allocate scarce resources. It is therefore to be expected that shortages could have significant impacts on consumers, businesses, and the overall functioning of the economy. Yet, despite their importance, there has been limited research on the long-term trends and patterns of shortages across various sectors.

In the first part of this paper, we construct a monthly shortage index for the United States, using newspaper articles from 1900 to 2023. The shortage index is a news-based indicator of the intensity of shortages of labor, materials, goods, and energy in the United States. To construct the index, we use news articles from six major U.S. newspapers, amounting to about 20,000 articles per month (approximately 25 million articles over the entire sample). The index covers a wide swath of history (both domestic and global events), and is much higher during periods of increased economic turmoil, such as the World Wars and the 1970s oil crises. It spikes considerably during the COVID-19 pandemic, reaching its highest level in the last 40 years, however, there are some other past spikes of comparable or larger size. We validate that our index is plausible with a narrative investigation and an AI-based text analysis.

In the second part of this paper, we show that shortages lead to persistently higher inflation and lower economic activity, using predictive regressions, a structural vector autoregressive (VAR) model, and a simple forecasting exercise.

We start with a series of predictive regressions in which we estimate the effect of shortages on measures of inflation and real growth. We find that shortages place downward pressure on GDP and upward pressure on prices, particularly in durable consumption and private fixed investment. Our results are robust to the inclusion of variables that are associated with inflation—for instance commodity prices and wages—indicating that our shortage index provides additional information not captured by traditional indicators.

To examine the variation in the effects of shortages over time, we estimate our regressions on a rolling sample. We find that shortages have consistently contributed to higher inflation over time. The effects of shortages on economic growth has been modest but negative for most of the sample, with two periods of higher growth associated with shortages: World War II and the COVID-19 pandemic. We interpret these results as a sign that both supply and demand

forces are relevant for analyzing shortages.

To further investigate the interplay between shortages and the business cycle, we estimate a structural vector autoregressive (VAR) model. In the model, we identify “traditional” demand and supply shocks, and shocks to shortages capturing, for instance, atypical market adjustment in response to sudden shifts in economic conditions, to regulation—such as mandated price ceilings or quantity rationing—, or weather and geopolitical events that impede the regular flow of goods in an economy. The advantage of a fully identified model compared to the predictive regressions is that it leads to an interpretation of movements in GDP, inflation, and shortages that is comprehensive and economically palatable.

We find that not all shortages are created equal. An important part of shortages in the 1950s are driven by demand shocks, associated with the Korean war, pent-up demand after World War II, and the rapid economic recovery that accompanied the transitioning from a wartime to a peacetime economy. The oil embargo of OPEC producers caused an unusual rise in shortages in the 1970s, over and above the direct effect of commodity price shocks. During the Covid pandemic, shortages emerged as a confluence of demand, supply, and shortage shocks, causing a persistent increase in inflation.

Our approach builds on previous work using news-based measures to track economic phenomena, such as the Economic Policy Uncertainty Index ([Baker et al., 2016](#)) and the Geopolitical Risk Index ([Caldara and Iacoviello, 2022](#)). However, to our knowledge, this is the first attempt to create a comprehensive shortage index for the United States spanning over a century. By examining the evolution of the index over a long period of time in conjunction with other economic indicators, we can gain insights into the causes and consequences of shortages and inform policy responses.

A number of other studies have also used news sources to construct indicators of shortages such as ours. As early as 1997, [Lamont \(1997\)](#) created a hand-coded news-based indicator of shortages using Wall Street Journal headlines. This was followed by additional attempts to measure shortages or related concepts via news-based indices ([Chen and Houle \(2023\)](#) for Canada, [Burriel et al. \(2023\)](#) for several advanced economies), as well as a more direct measure of supply chain pressures based on a variety of factors such as transportation costs ([Benigno et al., 2022](#)). Recently, [Pitschner \(2022\)](#) and [Bernanke and Blanchard \(2023\)](#) have tackled the intersection of shortages and the behavior of inflation during COVID-19 pandemic. In particular, [Bernanke and Blanchard \(2023\)](#), who use Google Trends to identify shortages, lay

down a model which seeks to explain the causes of pandemic-era inflation, and contend that shortages have a “strong but temporary effect” on inflation.

Relative to this literature, we make two main contributions. Our index is the first comprehensive measure of shortages spanning over 125 years, covering periods as diverse as the World Wars, the oil shocks of the 1970s and the 1980s, and the COVID-19 pandemic. Furthermore, we use both univariate regressions and a structural VAR analysis and find that shortages have persistent effects on inflation, more so than documented by some studies. Finally, we also find that while shortages have generally reflected a mix of demand and supply shocks, their realized “all-in” effects have been more in line with those of supply shocks.

The remainder of the paper is structured as follows. Section 2 discusses the construction of the index, presents it, and discusses its evolution over time. Section 3 validates the index. Sections 4 and 5 explore the relationship between the shortage index and economic activity and inflation in the United States, using predictive regressions and a VAR, respectively. Section 6 briefly illustrates the role of the shortage index as a predictor of inflation. Section 7 concludes and discusses potential applications and future research directions.

## 2 The Shortage Index

In this section, we first discuss the construction of the shortage index. We then show how the index captures key episodes of shortages in U.S. history.

### 2.1 Construction of the Index

The shortage index is a monthly news-based indicator that measures the intensity of shortages of materials, goods, labor, and energy in the United States. It is constructed from a sample of approximately 20,000 news articles per month, spanning the period from 1900 through the end of 2023—encompassing about 25 million articles. These articles were published in major U.S. newspapers, namely the *Boston Globe*, the *Chicago Tribune*, the *Los Angeles Times*, *The New York Times*, *The Wall Street Journal*, and *The Washington Post*. Each month, the index counts the number of articles that discuss energy, food, industry or labor shortages—the set  $\mathcal{S}$  depicted in Figure 1—normalized by the total number of articles, denoted as the set  $\mathcal{A}$ . A higher index value reflects a greater intensity of shortages. Below, we outline the steps that led to the creation of the search query used to isolate set  $\mathcal{S}$  and reported in Table 1.

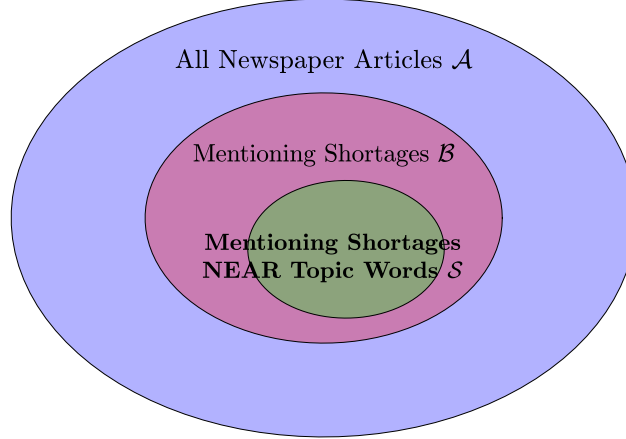


Figure 1: Grouping of Newspaper Articles for the Construction of the Shortages Index.

We first construct a broader set of articles,  $\mathcal{B}$ , represented by the pink area in Figure 1. Articles in this set must include at least one mention of a shortage term—namely ‘shortage’, ‘scarcity’, ‘bottleneck’, or ‘rationing’—in combination with an economics-related term, such as ‘economy’, ‘market’, or ‘commerce’. The shortage terms used are those more frequently associated with economically relevant constraints on production capacity of a country or on the availability of goods to consumers.<sup>1</sup> The inclusion of an economics term in the search query helps reduce the likelihood of false positives—articles that mention shortages unrelated to the economic phenomenon being measured.

Next, we draw a random sample of just over 3,300 articles from the set  $\mathcal{B}$ . From these articles, we construct a list of the 1,000 most frequent collocates, words that appear within five words of the shortage terms. We then select collocates that are indicative of shortages in specific sectors of the economy or goods. The most common collocates, excluding stopwords, include ‘oil’, ‘water’, ‘war’, ‘time’, ‘coal’, ‘days’, ‘food’, ‘cars’, ‘people’, ‘government’, ‘million’, ‘labor’, ‘state’, ‘home’, ‘steel’, and ‘fuel’. Words like ‘time’, ‘days’, ‘people’, and ‘government’, which convey little substantive information about shortages, are subsequently removed. The remaining collocates are grouped into four topics: food, industry, labor, and energy.

With these three lists of terms—shortage terms, economics terms, and topic-specific terms, we construct the search query shown in Table 1. An article is included in the shortages index—the set  $\mathcal{S}$  in Figure 1—if two conditions are met: first, a shortage term must appear

<sup>1</sup> Potential synonyms of shortage, such as ‘lack’ or ‘paucity’ or ‘insufficiency’ were excluded, as they have a broader range of meanings and are less likely to be specifically associated with economic shortages.

within five words ( $N/5$ ) of a topic-specific term; second, the article must contain at least one economics term. If an article meets the first condition for topic-specific terms from more than one category, it is counted once for each category. Consequently, the total number of shortage articles in any given period is the sum of articles from each category. This method allows us to give greater weight to articles discussing multiple types of shortages, enhancing the overall index’s sensitivity. In Section 3, we demonstrate that requiring proximity between shortage terms and topic-specific terms is crucial for reducing false positives and improving search accuracy.

Our classification of articles into four topics is further supported by the Latent Dirichlet Allocation (LDA) analysis, performed ex-post on a sample of articles that meet our inclusion criteria. LDA is a widely-used, unsupervised machine learning technique for topic modeling in natural language processing. It identifies hidden topics within a document corpus by analyzing word co-occurrence. The algorithm has two key inputs: the first is the text corpus, which in this case consists of a random sample of 13,623 newspaper article abstracts that mention shortages and meet our index inclusion criteria. This sample represents about 4 percent of the approximately 330,000 articles included in our index from 1900 to 2023.<sup>2</sup> The second input is the number of topics, which we set to four.

The results of the LDA analysis are presented in Figure 2. The most frequent terms for each topic are displayed as word clouds with the following topic categorizations: Topic 1 corresponds to energy, Topic 2 to water, food, and agricultural products, Topic 3 to industrial products (e.g., coal, steel, railroads, cars), and Topic 4 to jobs. The stacked bar chart at the bottom of the figure visualizes the topic mixture for each abstract—13,623 bars in total—sorted by year. Each year’s news coverage reflects varying degrees of focus on different shortage topics. In the early part of the sample, discussions are primarily centered around industry-related shortages, with food and water shortages receiving some attention. Energy shortages become particularly prominent during the 1970s, largely due to events like the oil crisis, which significantly impacted the global economy. In the more recent period, following the COVID-19 pandemic, labor shortages have increasingly dominated the news, reflecting the widespread disruptions in workforce availability and supply chains that have persisted post-pandemic.

---

<sup>2</sup> Following standard practice, we pre-process the text by removing stopwords, numbers, and reducing words to their root forms (stemming).

## 2.2 Shortages In History

We now present the shortage index, examine its spikes, and consider the historical context in which these spikes occurred. Figure 3 plots the shortage index at a monthly frequency from 1900 through the end of 2023. The index is calculated by taking the monthly share of articles discussing economic shortages ( $\mathcal{S}/\mathcal{A}$ ) and scaling it so that it averages 100 over the period 1900-2023.<sup>3</sup> In Table 2, we list the thirty largest spikes in the index accompanied by a description of the key events that correspond to each episode.<sup>4</sup>

The index shows considerable variation over time, with the most significant spikes linked to events connected to the four classification topics described above: energy, food, industry, and labor. A breakdown of the index by topic is presented in Figure 4.

Geopolitical events, especially wars, are closely associated with severe shortages. For example, the index rises dramatically during World War I and surges again during World War II, peaking at over 1,000 (ten times the sample mean) in January 1943, with spikes recorded across all shortage categories. Other geopolitical events, such as the Suez Crisis and the Iraqi invasion of Kuwait in 1990, also coincide with substantial spikes. Many of these spikes are tied to energy shortages, often resulting from wars and instability in the Middle East. The oil shocks of the 1970s, which led to the only other instance where the index exceeded 1,000, are a key example of this pattern. Labor shortages, particularly those caused by strikes, have historically been significant, especially in the early part of the sample. For instance, coal-related strikes in 1903, 1919, and 1922 caused notable spikes in the index. However, strike-related shortages become less frequent in more recent decades, as labor markets evolved and other factors began to dominate the landscape.

More recently, the index has increased multiple times during the COVID-19 pandemic. The first spike corresponds to shortages in medical equipment and healthcare workers at the onset of the pandemic. A larger spike occurred at the beginning of 2022, driven by global supply bottlenecks as economies reopened after prolonged mobility restrictions in 2020 and 2021. As seen in Figure 4, the second spike primarily reflected shortages in labor and industry

---

<sup>3</sup> Specifically, the share is divided by its mean over the 1900–2023 period and then multiplied by 100.

<sup>4</sup> To calculate the spikes, we first extract the residuals from a regression of the shortage index,  $h_t$ , on its values two months before and two months after. The largest residuals form our list of candidate spikes. We then discard any spike if it occurs in a month,  $t$ , where the index is not a local maximum within the 13-month window  $[t - 6, t + 6]$ . This second step ensures that we are capturing true local peaks. From the remaining observations, we report the thirty largest residuals in Table 2.

components.

Finally, Figure 5 provides a breakdown of the shortage index into a US-specific component and a foreign component. The foreign component isolates articles that meet the original search criteria and include the mention of an international term, such as the name of a foreign country or major foreign city. As shown in the figure, the US-specific and foreign components move in lockstep, with peak values generally occurring around the same time. This finding suggests that the events captured by our index are not confined to the US but are global in scope.

### 3 Assessing the Accuracy of the Shortage Index

We conduct two separate exercises to assess the accuracy of the index. First, we verify that the newspaper articles included in the index indeed mention concerns related to shortages. Second, we evaluate whether the index aligns with alternative measures or proxies for shortages during the limited time periods where these alternatives exist and overlap with our data.

#### 3.1 Validation of the Shortage Index

We verify that our index accurately measures shortages by minimizing Type-I and Type-II errors. To do this, we sample the abstracts of 872 articles from the shortage set  $\mathcal{S}$ .<sup>5</sup> By construction, each of these articles contains at least one business-related term as well as a mention of scarcities, shortages, or bottlenecks in proximity to a topic word such as energy, food, industry, or labor. For each article, we extract the first snippet of text that references shortages. These snippets are centered around the shortage word and limited to 110 characters, drawing inspiration from Twitter’s original 140-character limit. This length strikes a balance between brevity and reducing computational and cognitive load. For example, two snippets from our sample include:

- “Although demand remains strong... the resulting supply shortage of German manufacturing goods could also...” (2021)

---

<sup>5</sup> Our search query for calculating the index is based on six U.S. newspapers listed in Section 2. Due to technical reasons, abstracts have not been available for these newspapers since 2015. For the period 2015–2023, we sampled abstracts from a broader set of newspapers, including some based in the U.K., Canada, and Australia. A detailed list of the 872 abstracts and their sources used for validation is available upon request. Abstracts are typically short portions of text, often containing the opening sentences or the first paragraph of the article.



- “...men interested in the industries affected by the shortage of steel are anxious to see the strike settled.” (1901)

We then used the Claude AI assistant ([Anthropic, 2024](#)) to determine whether each snippet mentions current or prospective shortages, rationing, scarcity, or bottlenecks related to goods, labor, materials, food, or water. Claude was instructed to return a table of results, coding articles as 1 (shortage mentioned), 0 (shortage not mentioned), or 99 (unsure whether the existence of shortages was mentioned). In addition to the classification, Claude was asked to provide a brief explanation for each snippet’s coding.

Before initiating the classification, we provided Claude with examples, listed in the Appendix, of how we would code the snippets and ensured that the training sample included false positives, such as mentions of the lack or end of shortages. Use of AI for validation is not foolproof, but we found Claude did as good a job as a human, for instance by extrapolating the context of a particular sentence to a particular country or person. For instance, for the sentence “economy may be slowing but Lowe is banking on labour shortages gradually leading to an increase...”, Claude classified the text as 1 and added that “Reserve Bank [of Australia] expecting labour shortages to lead to wage growth.”

The results of the audit are in Table 3. Out of 872 articles belonging to the set  $\mathcal{S}$ , only 6.3 percent were deemed by ClaudeAI as false positives. For instance, Claude classified the snippet “...says that while there is not necessarily a shortage of people wanting to work in management” as 0 with the explanation “No shortage of people wanting to work in management”. Additional examples are available in the Appendix.

We then repeat the audit for a sample of 298 articles not belonging to the set  $\mathcal{S}$ . Out of these 298 articles, only one appears to mention shortages but is not captured by our search query (“recycling of newsprint was held back by a shortage of deinking plants”). Of note, our search query deliberately did not include the word “plants” since in preliminary attempts we found instances of false positives associated with this word.<sup>6</sup>

Finally, we confirm that restricting the search to include shortage words in proximity to terms related to goods, labor, food, or energy significantly improves the accuracy of the search. In the broader set of articles that allows, but does not require, the presence of shortage words near these key terms, 84.2 percent mention economically relevant shortages. This results in a

---

<sup>6</sup> See for instance the article “Brighten Up Indoors With Colorful Plants” (Los Angeles Times, Feb. 4, 1996) that states “there’s no shortage of plants with brightly colored foliage to liven up your kitchen.”

Type I error of 15.8 percent, which is considerably higher than in our preferred search query. False positives in this broader set, which are not captured by our preferred method, include articles referencing non-economic shortages, such as shortages of political campaign funds, lack of good baseball photos, legislative bottlenecks, and even shortages of sunshine.

### 3.2 Comparison with Other Indicators of Supply Constraints

In this section, we compare our index with related measures of supply constraints.

Figure 6 plots our shortage index alongside the New York Fed Global Supply Chain Pressure (GSCPI) Index, the Supplier Delivery Index (SDI), and the [Burriel et al. \(2023\)](#)’s Supply Bottlenecks Index (SBI) for the US. The GSCPI, published by the Federal Reserve Bank of New York, measures global supply chain conditions using data on both manufacturing and transportation costs. The SDI, published by the Institute of Supply Management (ISM), is based on a monthly survey asking firms whether they are experiencing longer or shorter wait times compared to the previous month. The SDI value represents the share of respondents reporting longer wait times, plus half the share reporting no change. The SBI, produced by [Burriel et al. \(2023\)](#), uses a text-based newspaper search to quantify supply chain issues. To facilitate the comparison, we standardize each of the variables to have a mean of 0 and a variance of 1 over the period in which they overlap.

Figure 6 shows that our index shares similar features to these three indicators. Over the entire existence of the GSCPI, the correlation between our index and the GSCPI is 0.73. Both measures increase sharply at the onset of COVID-19 in early 2020, and again in the beginning of 2022 as supply chain bottlenecks took hold. The correlation between our index and the SDI is 0.25. Nonetheless, both measures spike around the 1979 oil crisis and COVID-19. One possible explanation is that both events caused significant transportation delays, whether due to rising fuel costs or supply bottlenecks, which may have contributed to shortages experienced by manufacturers. Finally, our index has a high correlation of 0.90 with the U.S. SBI, indicating a strong alignment between these two measures over the common sample.

Thus, while all indexes contain useful information to analyze the extent of shortages and broader supply constraints, a key advantage of our index relative to these indicators is that our index is available over a much longer period of time, thus being particularly amenable to historical applications for research.

## 4 Shortages as Predictors of Inflation and Activity

Figure 7 illustrates the positive relationship between U.S. inflation and our indicator of shortages for a sample starting in 1940.

In this section, we formally explore the relationship between shortages, near-term inflation, and economic activity using a simple regression framework. Specifically, we estimate the following predictive regression:

$$\Delta Y_{t+h} = \alpha + \beta SHORTAGE_t + \sum_{i=0}^p \mathbf{X}_{t-p} + \varepsilon_{t+h}, \quad (1)$$

where  $\Delta Y_{t+h} = \frac{400}{h} \ln(\frac{Y_{t+h}}{Y_t})$  is the annualized log change of a variable of interest  $Y_t$  between period  $t$  and forecast horizon  $h$  and  $SHORTAGE_t$  denotes the level of the shortage index.  $\mathbf{X}$  is a vector of control variables.

We use quarterly data from 1950 through 2023 from the National Income and Product Accounts (NIPA) for real per-capita GDP, personal consumption expenditures, and private fixed investment. Inflation for each category is measured using the associated price deflator. For GDP and its price deflator, we extend the sample back to 1900 using data from [Ramey and Zubairy \(2018\)](#). Data on the total population also come from [Ramey and Zubairy \(2018\)](#), which we extend through 2023 using the POP series from FRED.

For each price and economic activity indicator, we estimate regression (1) separately by OLS. As control variables, we include quarterly changes of the dependent variable and the associated economic indicator or price deflator, both contemporaneously and with three lags. For instance, the predictive regression for real GDP growth includes contemporaneous and lagged values of both real GDP growth and inflation, measured by the log change in the GDP deflator.

The results for the four-quarter-ahead predictive regressions are presented in Table 4, while results for the one- and eight-quarter horizons are reported in Tables A.1 and A.2, respectively. To facilitate comparison across variables, we report standardized estimates of the coefficients  $\beta$ . A standardized coefficient represents the movement of the dependent variable (in standard deviation units) in response to a one standard deviation change in the explanatory variable.

The first two columns of Table 4 report estimates for the full sample. An increase in shortages is associated with a rise in inflation (first column) and a decline in economic activity

(second column), the typical effects of supply-side disruptions. The inflationary effects of shortages are fairly evenly distributed across GDP components, with prices in the services consumption category being the least impacted. While durable goods consumption and private fixed investment decline, there is no statistically significant effect on the consumption of non-durable goods or services.

To quantify the economic effects of shortages, take for instance the standardized coefficients estimated for durable consumption. A one-standard deviation increase in the shortage index is associated with an increase in durable goods inflation of 0.75 percentage point (0.25 standard deviation), and with a decline in durable goods consumption growth of 1.75 percentage points (-0.25 standard deviation). The coefficients for GDP imply an increase in inflation measured by the GDP deflator of 0.5 percentage point and a reduction in real GDP growth of 0.3 percentage point.

The remaining columns of Table 4 show that the results for the full sample hide notable time variation in the relationship between shortages, inflation, and economic growth. To illustrate this point, we split the sample in 2013:Q4, so as to better isolate the role of shortages in the 10 years since 2014 that include the global pandemic and the subsequent recovery. In the sample running from 1950 through 2014, the effects of shortages on inflation and activity are precisely estimated. In addition, the reduction in economic growth from an increase in shortages is widespread across sectors, sparing only private consumption of services. In the sample starting in 2015 and encompassing the COVID-19 pandemic, the effects of shortages on inflation are substantially larger, albeit less precisely estimated at the sectoral level. The coefficient of shortages on economic activity becomes positive, but is not statistically significant. We interpret these results as suggestive evidence that demand forces may have played a more prominent role in driving shortages during and after the COVID-19 pandemic, compared to the pre-COVID period.

A potential concern with these results is whether shortages can provide information beyond what is already embedded in other macroeconomic variables. For example, oil prices may move similarly to the energy component of our index. In Table 5, we test this possibility by adding in several potential variables: oil prices, commodity prices, wage growth, and inflation expectations. The top row of Table 5 reports the same baseline GDP estimate seen in the top row of Table 4. Each of the next rows adds the named variable to the baseline as an additional

control.<sup>7</sup> We find that our shortage index is informative, as its coefficient typically remains significant. Columns 1 and 2 show that the effect of shortages is somewhat attenuated but still similar in magnitude to the effects seen in Table 4. In the next four columns, we again partition the sample into two periods. As in Table 4, the pre-COVID effects are stable and highly significant. The effects in the later part of the sample are less significant but remain largely in line with the baseline results. Overall, these results suggest that our index encodes additional information over and above what is in traditional macroeconomic variables.

One might wonder whether the COVID-19 experience has been unique, or whether other historical episodes of widespread shortages were also associated with a strong role for demand in driving movements in shortages. We investigate this question by estimating the specification in equation (1) on a rolling sample using a 30-year window. For this analysis, we limit our attention to real GDP growth and to inflation, as measured by the GDP deflator.<sup>8</sup> Figure 8 shows the results. Starting with the sample including World War II, shortages have consistently exerted inflationary effects. The 10 years prior to the pandemic are a notable exception, and the effect of shortages on inflation become negligible as well as imprecisely measured. Shortages are associated with inflationary effects when the sample includes the years around the pandemic.

Shortages have exerted economically meaningful and statistically significant adverse effects on economic activity starting in the 1970s up until the COVID-19 pandemic. In contrast, an increase in our index is associated with higher activity around just two historical episodes: the pandemic and World War II. Rationing, which was commonplace during the war, implies that demand outstripped supply. Thus both episodes, through their own unique circumstances, were times in which high demand propped up economic activity. Furthermore, the imposition of price caps during the war may have contributed to keeping inflation more muted than expected.

In the next section, we take a complementary approach in investigating how shortages interact with demand and supply factors, estimating a structural VAR model.

---

<sup>7</sup> For details, refer to the notes for Table 5.

<sup>8</sup> The use of the GDP deflator allows extending the sample all the way back to 1900.

## 5 Shortages and Activity: A VAR Analysis

In this section, we present a structural VAR model of the U.S. economy that incorporates shortages. The primary advantage of a full-blown empirical model discussed in Section 4 is that it provides a more comprehensive explanation of how movements in shortages interact with broader economic activity. Another advantage of the VAR model is that it can account for time-varying reduced-form effects of shortages on economic activity to the extent that the relative importance of different factors simultaneously affecting shortages and economic activity changes over time.

### 5.1 The Model

The model consists of five indicators measured at quarterly frequency: economic activity, inflation, commodity prices, shortages, and the short-term interest rate, summarized by the vector  $\mathbf{X}_t = (y_t, \pi_t, c_t, h_t, r_t)'$ . Economic activity is measured as the four-quarter percent change in real GDP. Inflation is the four-quarter percent change in the headline CPI index, commodity prices are the four-quarter change percent change in the Reuters–CRB (Commodity Research Bureau) Spot Commodity Price Index for raw industrials and foodstuffs. The short-term interest rate is the annualized 3-month treasury bill rate, and shortages are expressed in levels (standardized). All series are demeaned.

The vector autoregressive representation of the economy is given by:

$$\mathbf{A}\mathbf{X}_t = \sum_{j=1}^p \mathbf{B}_j \mathbf{X}_{t-j} + \mathbf{C}\mathbf{u}_t, \quad (2)$$

where  $\mathbf{u}_t = (u^S, u^D, u^C, u^H, u^R)'$  is a vector of structural shocks,  $p$  is the lag length, and  $\mathbf{A}$ ,  $\mathbf{B}_j$  for  $j = 1, \dots, p$  and  $\mathbf{C}$  are matrices of structural parameters. The structural shocks have zero mean and a diagonal variance–covariance matrix  $E[\mathbf{u}_t \mathbf{u}_t'] = \Sigma_u$ . Without loss of generality, we normalize one entry on each row of  $\mathbf{A}$  to 1.

The following equations describe the structural relationships among variables:

$$\pi_t = u_t^S + \kappa y_t + [\mathbf{b}^S]' \mathbf{z}_{t-1}, \quad (3)$$

$$y_t = u_t^D - \delta \pi_t + [\mathbf{b}^D]' \mathbf{z}_{t-1}, \quad (4)$$

$$c_t = u_t^C + \chi_D u_t^D + \chi_S u_t^S + [\mathbf{b}^C]' \mathbf{z}_{t-1}, \quad (5)$$

$$h_t = u_t^H + \theta_D u_t^D + \theta_S u_t^S + \theta_C u_t^C + [\mathbf{b}^H]' \mathbf{z}_{t-1} \quad (6)$$

$$r_t = u_t^R + \alpha_H h_t + \alpha_Y y_t + \alpha_\pi \pi_t + \alpha_C c_t + [\mathbf{b}^R]' \mathbf{z}_{t-1}, \quad (7)$$

where  $\mathbf{z}_{t-1} = (\mathbf{X}'_{t-1}, \mathbf{X}'_{t-2}, \dots, \mathbf{X}'_{t-p})'$ . Equation (3) describes an inverse aggregate supply equation, where inflation is positively related to output ( $\kappa > 0$ ) and subject to adverse supply shocks  $u^S$ . Equation (4) represents aggregate demand, with demand  $y$  negatively related to inflation ( $\delta > 0$ ) and affected by demand shocks  $u^D$ . Equation (5) models commodity prices, which respond to demand and supply shocks alongside commodity-specific shocks.

Equation (6) describes the relationship between shortages and the remaining block of the model. We assume that shortages reflect “regular” business cycle movements caused by supply and demand shocks, shocks in commodity markets, and an “exogenous” component that proxies for newsworthy disruptions to the regular flow of goods, services and factors of production in the economy. The shocks  $u^H$  capture “shortages shocks,” that is, any unusual combination of unpredictable events that cause demand to temporarily exceed supply.<sup>9</sup> Finally, equation (7) describes a monetary policy rule, with the term  $u^R$  denoting monetary policy shocks.

Examples of exogenous shortages may include atypical market adjustments in response to sudden shifts in demand and supply. For instance, demand reallocation from one sector to another can create bottlenecks that require temporary rationing. Other examples include regulatory shocks, such as mandated price ceilings or quantity rationing, which disrupt the normal functioning of a market economy; panic buying that leads to rationing of certain goods when social norms prevent large price adjustments; and extreme weather events or sudden

---

<sup>9</sup> Our assumption is that shortages reflect, within the quarter, movements in economic conditions caused by traditional demand and supply shocks. This assumption differentiates our analysis of the effect of shortages on activity and inflation from the work of [Burriel et al. \(2023\)](#), who order shortages first in a monthly VAR from 1990 through 2020; and from the work of [Bernanke and Blanchard \(2023\)](#), who use a quarterly VAR from 1990 through 2023 and also assume that shortages affect inflation within the quarter.

geopolitical shocks that impede the smooth flow of goods within an economy.<sup>10</sup>

According to our timing specification, exogenous shortage shocks cannot directly impact activity, prices or commodity prices within a quarter (while monetary policy is allowed to respond to shortages within the quarter). Accordingly, the effects of shortages shocks manifest with at least a one-quarter delay, and depend only on the lagged feedback from shortages to GDP, inflation and commodity prices captured by the coefficients  $\mathbf{B}_j$  for  $j = 1, \dots, p$ .

The matrices summarizing the system are:

$$\mathbf{A} = \begin{bmatrix} 1 & -\kappa & 0 & 0 & 0 \\ 1 & \delta & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ -\alpha_\pi & -\alpha_Y & -\alpha_C & -\alpha_H & 1 \end{bmatrix}; \quad \mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \chi_S & \chi_D & 1 & 0 & 0 \\ \theta_S & \theta_D & \theta_C & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad (8)$$

with the impact matrix relating variables at time  $t$  to structural shocks given by:

$$\mathbf{D} \equiv \mathbf{A}^{-1}\mathbf{C} = \begin{bmatrix} \frac{1}{1+\kappa\delta} & \frac{\kappa}{1+\kappa\delta} & 0 & 0 & 0 \\ \frac{-\delta}{1+\kappa\delta} & \frac{1}{1+\kappa\delta} & 0 & 0 & 0 \\ \chi_S & \chi_D & 1 & 0 & 0 \\ \theta_S & \theta_D & \theta_C & 1 & 0 \\ \gamma_S & \gamma_D & \gamma_C & \alpha_H & 1 \end{bmatrix}, \quad (9)$$

where the reduced-form parameters describing the interest rate response to shocks are given by  $\gamma_S = \frac{\alpha_\pi - \delta\alpha_Y}{1+\kappa\delta} + \alpha_C\chi_S + \alpha_H\theta_S$ ,  $\gamma_D = \frac{\alpha_Y + \kappa\alpha_\pi}{1+\kappa\delta} + \alpha_C\chi_D + \alpha_H\theta_D$ , and  $\gamma_C = \alpha_C + \theta_C\alpha_H$ .

## 5.2 Priors and Estimation

Following [Baumeister and Hamilton \(2019\)](#), we view identification and estimation of this structural VAR model as a special case of a Bayesian VAR with prior restrictions imposed on the parameters of the matrices  $\mathbf{A}$ ,  $\mathbf{B}_j$  and  $\mathbf{C}$ . In turn, these prior restrictions may aid identification of the structural shocks and their effects. To see why, consider the example in which one assumes a flat within-quarter Phillips curve, so that  $\kappa = 0$ . The matrix  $\mathbf{D} = \mathbf{A}^{-1}\mathbf{C}$  relating variables to structural shocks becomes lower triangular and the VAR model has a

---

<sup>10</sup> [Kahneman et al. \(1986\)](#) describes how fairness considerations may lead to shortages and inefficient market functioning in the presence of unexpected shifts in demand or in supply.



familiar recursive interpretation with variables ordered as  $(y_t, \pi_t, c_t, h_t, r_t)$ . Our strategy is to estimate directly the parameters of the matrices  $\mathbf{A}$ ,  $\mathbf{B}_j$  and  $\mathbf{C}_j$  by imposing prior distributions.<sup>11</sup> In estimating the model, we set the number of lags  $p$  at two.<sup>12</sup>

Prior distributions for the key parameters are plotted in red in Figure 9. The gray areas in Figure A.1 plots impulse responses at the prior mean. As the figure illustrates, we view the prior distributions as imposing no more than reasonable bounds on the shape and magnitude of the impulse responses. Of note, the confidence intervals for the impulse responses at the prior are much wider than the corresponding posterior intervals, an indication that the model is informative about the model parameters.

- The priors for  $\kappa$  and  $\delta$  in the supply and demand equations (3) and (4) are inverse gamma with mean 0.2 and standard deviation 1. At the prior mean, the slope of the aggregate supply curve implies that inflation rises on impact by 0.2 percentage points in response to a demand-driven, 1 percent increase in activity. By the same token, the slope of the aggregate demand curve implies that, on impact, a 1 percent supply-driven increase in inflation reduces GDP growth by 0.2 percentage point. These prior assumptions—that we relax later—are broadly in line with existing studies including evidence from estimated DSGE models of the U.S. economy: see for instance [Smets and Wouters \(2007\)](#).
- The priors for  $\chi_S$  and  $\chi_D$  in the commodity price equation are drawn from an inverse gamma distribution with mean equal to 2 and standard deviation equal to 5. Thus, the prior imposes that commodity prices rise and react strongly—but with a very generous range of possible outcomes—to both expansionary demand shocks and contractionary supply shocks.
- The priors for  $\theta_S$ ,  $\theta_D$ ,  $\theta_C$ —the parameters dictating how shortages response to macroeconomic disturbances—are drawn from an inverse gamma distribution with mean 0.2 and standard deviation 1. This choice of priors embeds the assumption that shortages and associated news can reflect standard demand and supply shocks, as well as shocks to commodity markets, within a quarter. Of note, embedded in our choice of priors is the assumption that expansionary demand shocks are associated with an increase in

---

<sup>11</sup> This strategy is different from the standard approach of exactly identified VAR models where one sets enough restrictions on the parameters of the model to uniquely identify all shocks.

<sup>12</sup> To avoid overfitting the model, the set the cross-partial terms beyond the first lag at zero. Table A.3 reports the prior distributions and the estimated parameters of the model.

shortages, that expansionary supply shocks are associated with a decrease in shortages, and that shocks that lead to higher commodity prices are associated with an increase in shortages (that is, the support of  $\theta_S$ ,  $\theta_D$  and  $\theta_C$  is positive).

- The priors for the monetary policy rule equation assume that there is a systematic positive response of interest rates to inflation and GDP, with a short-run response coefficients of 0.1; the long-run response is assumed to be larger due to the autoregressive parameters in the rule itself, as captured by the coefficients  $\rho_{1,RR}$  and  $\rho_{2,RR}$ .

The support of the prior densities is wide, encompassing substantially smaller and larger elasticity values relative to the mean. As discussed in the empirical results, the data and the model structure lead to revising both in the location and spread of these distributions.

Finally, the prior mean for the coefficients in the matrices  $\mathbf{B}_j$ , for  $j = 1, 2$ , is set so that variables have first and second-order autocorrelation of 0.65 and 0.10, respectively. All lagged indirect effects have a prior mean of zero, with the exception of the lagged responses of inflation and activity to interest rate shocks, which are assumed to be negative, and of the lagged response of inflation to commodity price shocks, which embeds a small amount of pass-through.

## 5.3 Results

We estimate the model on quarterly data from 1950:Q1 through 2023:Q4 using the software Dynare. The first 20 observations are used as a training sample. To construct the estimates of interest, we take 20,000 draws from the posterior distribution using a Random Walk Metropolis-Hastings algorithm.

### Parameter Estimates

Figure 9 shows the prior distributions and the estimated posterior distributions for some of the key parameters of the model. The location and spread of all posterior distributions is substantially updated compared to the priors, revealing that the data and the structure of the model are informative about the structural parameters.

The posterior median for the parameter  $\delta$  in the demand equation is around 0.11, which implies that a supply shock that increases inflation by 1 percentage point leads to a decrease in GDP of about 0.11 percentage point. The posterior mean for the parameter  $\kappa$  is 0.11—which

implies that a demand shock that raises GDP growth by 1 percentage point leads on impact to an increase in inflation of around 0.11 percentage point. All told, the posterior distributions reveal a “flatter” supply curve (the posterior mean of  $\kappa$  is lower than the prior mean) and a “steeper” demand curve (the posterior estimate of  $\delta$  is lower than the prior mean).

The top right panel of Figure 9 illustrates one of the underlying forces driving the inflationary effects of shortages in our model. After controlling for current and lagged movements in all other variables, the posterior density measuring the effect of shortages in period  $t$  on inflation in period  $t + 1$ —as captured by the parameter  $\rho_{1,PH}$ —is positive, indicating a high likelihood of lagged shortages having a positive effect on inflation.

Estimates for  $\theta_D$ ,  $\theta_S$  and  $\theta_C$  are consistent with the notion that shortages respond to the business cycle and to movements in commodity prices, albeit to a smaller degree than assumed by the priors. The posterior estimate for both parameters is positive but relatively small.

The last row of Figure 9 describes the estimated response of interest rates to inflation and output, which is positive. Additionally, the parameter estimates suggest a relative modest response of interest rates to increases in shortages.

## Impulse Responses

Figure 10 reports the impulse responses to the estimated shocks in the model. The solid line depicts the posterior mean, while the shaded areas represent the 80 percent posterior credible sets. All shocks are one-standard deviation in size.

As shown in panel (a), an exogenous increase in shortages is associated with a slow and persistent decline in GDP, alongside a rise in inflation. Inflation peaks about two years after the shock and remains elevated thereafter. GDP growth declines persistently, and there is an initial small decline in interest rates.<sup>13</sup>

Shortages can also arise endogenously in response to other shocks. Specifically, an expansion in aggregate demand and a contraction in aggregate supply result in higher shortages, as illustrated in panels (b) and (c). Similarly, a commodity-driven price shock, like those seen in the 1970s and 1980s and depicted in panel (d), also leads to increased shortages. Therefore, in our model, causality runs both ways: exogenous shocks to shortages can cause business cycle

---

<sup>13</sup> Incidentally, when we compare our analysis to [Bernanke and Blanchard \(2023\)](#), we find a significant and much more prolonged response of inflation to a shock to shortages. Using a Google-trends based measure of shortages starting in 2004, [Bernanke and Blanchard](#) report that the effects of shortages on inflation is small and short-lived, with the effects dying out after only one quarter.

fluctuations; exogenous shocks to demand and supply can trigger movements in shortages.

With all shocks scaled to one standard deviation, we can compare their relative importance for output and inflation. An important result is that, for a given shortfall in activity, the rise in inflation is smaller under a shortage shock than under a “traditional” contractionary supply shock, both for headline inflation and for commodity price inflation. In other words, news about shortages are a likely signal of higher inflation, but with more subdued-than-usual economic activity. Of note, while interest rates tends to rise following adverse supply shocks, shortages shocks seem to exert very limited upward pressure on interest rates.

### **The Effects of Shortages throughout History**

Figure 11 shows the historical decomposition of the model. A large part of the shortages in the early 1950s are driven by demand shocks, associated with the Korean war, pent-up demand after World War II, and the rapid economic recovery that accompanied the transitioning from a wartime to a peacetime economy. The oil embargo of OPEC producers caused an unusual rise in shortages in the 1970s, over and above the direct effect of commodity price shocks.

Figure 12 zooms in on the role of shortage shocks between 2020 and 2023. The COVID and post-COVID years can be separated into three distinct periods. The first period, which runs from 2020:Q1 through 2021:Q1, witnessed both a reduction in aggregate demand and tighter supply conditions. Shortages, primarily driven by lack of workers and supply chain issues, started to build up, but were offset by weak demand. The second period runs from the second quarter of 2021 through the third quarter of 2022 and is characterized by unusual shortage shocks stemming from supply chains strains and tight labor market conditions, in a context of strong aggregate demand and rising commodity prices. Throughout this period, shortage shocks start exerting upward pressure on inflation. The third period runs from 2022:Q4 through the end of sample in 2023:Q4 and is characterized by a slow decline in shortages, which nonetheless remain above their pre-pandemic level. Importantly, as the estimated effect of shortage shocks on inflation is delayed and long-lived, past shortages continue to be a prominent driver of inflation all the way through the end of 2023.

The estimated model also allows shedding light on the combined role played by demand and supply factors in driving shortages, inflation and economic activity. Figure 13 illustrates that the model views the rise in inflation as predominantly driven by “supply” shocks that drive inflation and economic activity in opposite directions.

## Robustness Exercises

Figure 14 illustrates the response of inflation to a shortage shock and the role of shortages and other shocks in the post-2020 rise and fall in inflation in the baseline model and three alternative versions. The role of shortages in explaining inflation is substantial across specifications, except if we make the strong assumption that the demand curve is flat (panel c), fixing the slope of the short-run inverse demand curve at  $\delta = 4$ .

## 6 Can Shortages Forecast Inflation?

In the last section of the paper, we examine the efficacy of the shortage index in forecasting inflation at the 12-month horizon, building on the work of [Stock and Watson \(1999\)](#). Specifically, we estimate the following specification:

$$\pi_{t+12} = c + \beta(L) \pi_t + \gamma(L) x_t + \delta(L) h_t, \quad (10)$$

where  $\pi_{t+12}$  is one-year ahead, 12-month percent change in the headline CPI index,  $\pi_t$  is current 12-month CPI inflation and  $\mathbf{x}_t$  is a vector of economic variables including unemployment and the 12-month change in oil prices,  $h_t$  is the shortage index, and  $\beta(L)$ ,  $\gamma(L)$  and  $\delta(L)$  are polynomial operators in the lag  $L$ . We choose 12 lags of each dependent variable, a window of 30 years, and implement a series of rolling forecasts starting in 2000:M1 through 2023:M12, encompassing both the Great Moderation period and the recent economic turbulence associated with the COVID-19 pandemic and its aftermath.<sup>14</sup> This exercise hews closely to a real-time forecasting exercise in that at each point in time one-year ahead forecasts of inflation are only made using data available up to that point.

Our primary finding is that a model incorporating inflation expectations, unemployment, oil prices and inclusive of our shortage index outperforms the competing model that excludes the shortage measure, both in the 1990-2019 period and in the more limited time window from 2020 through 2023. For the 1990-2019 period, the root mean square error (RMSE) of the forecast that does not incorporate shortages is 1.33, about 10 percent higher of the

---

<sup>14</sup> The first prediction is made for 1990:M1 using the model estimated using data from 1960:M1 through 1989:M12. The window is then moved ahead, one month at the time. To reduce the outsized role played by big swings in oil prices and unemployment during the pandemic, we use 12-month lagged moving averages of oil prices and unemployment in the right-hand side.

RMSE of the model incorporating shortages, which equals 1.19. The model incorporating shortages delivers a relatively better performance also in the 2020-2023 period: its RMSE is 0.98, compared to 1.67 of the model without shortages. Figure 15 illustrates how the model with shortages predicts a sharper rise and a much slower decline of inflation throughout 2022 and 2023, thus explaining the relatively smaller forecast errors ex post.<sup>15</sup>

This finding suggests that shortages provide valuable additional information for inflation forecasting, beyond traditional predictors such as unemployment and commodity prices.

## 7 Conclusions

This paper introduces a new monthly, newspaper-based shortage index for the United States spanning more than a century, from 1900 through present. The index captures the intensity of shortages across various sectors of the economy, including labor, materials, goods, and energy. The index exhibits significant spikes during periods of heightened economic turmoil, such as the World Wars, the oil crises of the 1970s, and the COVID-19 pandemic.

Validation exercises confirm the accuracy of the index in measuring shortages and reveal strong correlations with other indicators of supply constraints. Notably, our index covers a substantially longer time horizon than existing alternatives, making it a valuable tool for historical analysis.

Predictive regressions demonstrate that increases in the shortage index are associated with persistently higher inflation and lower economic activity. These effects are particularly pronounced for durable goods consumption and private investment. Furthermore, the inflationary impact of shortages appears to be much stronger in the post-2015 period, which includes the COVID-19 pandemic.

The paper has also presented structural VAR analysis that decomposes shortage movements into the endogenous response to traditional business cycle shocks, on the one hand, and exogenous shortage shocks, on the other. The results indicate that movements in shortages cannot be simply explained by traditional demand, supply, commodity and monetary shocks, and that surprise innovations to shortages can be an important driver of inflation. Historically, the 1950s saw shortage movements primarily driven by demand factors, while the 1970s

---

<sup>15</sup> We obtain similar results when we replace the unemployment rate with more comprehensive business cycle indicators such as the Brave-Butters-Kelley Monthly GDP growth measure.

experienced unusual shortage shocks stemming from the OPEC oil embargo, beyond the direct effects of commodity prices. During the COVID-19 pandemic, shortages emerge as a combination of supply, demand, and exogenous factors, leading to a persistent inflationary impact.

The new shortage index developed in this paper provides a comprehensive and long-run perspective on the prevalence, drivers, and economic consequences of shortages in the United States over the past century. Our findings highlight the complex interplay of forces behind shortages and their tendency to boost inflation. This index serves as a valuable addition to the toolkit of policymakers and researchers, enabling a deeper understanding of the role of shortages in shaping macroeconomic outcomes.

# References

- ANTHROPIC (2024): “Claude,” Conversational AI model developed by Anthropic, version from March 30, 2024. Available at <https://www.anthropic.com>.
- BAKER, S. R., N. BLOOM, AND S. J. DAVIS (2016): “Measuring Economic Policy Uncertainty,” *The Quarterly Journal of Economics*, 131, 1593.
- BAUMEISTER, C. AND J. D. HAMILTON (2019): “Structural interpretation of vector autoregressions with incomplete identification: Revisiting the role of oil supply and demand shocks,” *American Economic Review*, 109, 1873–1910.
- BENIGNO, G., J. DI GIOVANNI, J. J. GROEN, AND A. I. NOBLE (2022): “The GSCPI: A new barometer of global supply chain pressures,” *FRB of New York Staff Report*.
- BERNANKE, B. AND O. BLANCHARD (2023): “What caused the US pandemic-era inflation?” *Peterson Institute for International Economics Working Paper*.
- BURRIEL, P., I. KATARYNIUK, C. MORENO PÉREZ, AND F. VIANI (2023): “A new supply bottlenecks index based on newspaper data,” *Banco de Espana Working Paper*.
- CALDARA, D. AND M. IACOVIELLO (2022): “Measuring geopolitical risk,” *American Economic Review*, 112, 1194–1225.
- CHEN, L. AND S. HOULE (2023): “Turning Words into Numbers: Measuring News Media Coverage of Shortages,” Tech. rep., Bank of Canada.
- KAHNEMAN, D., J. L. KNETSCH, R. THALER, ET AL. (1986): “Fairness as a constraint on profit seeking: Entitlements in the market,” *American economic review*, 76, 728–741.
- LAMONT, O. (1997): “Do” shortages” cause inflation?” in *Reducing Inflation: Motivation and Strategy*, University of Chicago Press, 281–306.
- NEWHEY, W. K. AND K. D. WEST (1987): “Hypothesis testing with efficient method of moments estimation,” *International Economic Review*, 777–787.
- PITSCHNER, S. (2022): “Supply chain disruptions and labor shortages: COVID in perspective,” *Economics Letters*, 221, 110895.



- RAMEY, V. A. AND S. ZUBAIRY (2018): “Government spending multipliers in good times and in bad: evidence from US historical data,” *Journal of Political Economy*, 126, 850–901.
- SMETS, F. AND R. WOUTERS (2007): “Shocks and frictions in US business cycles: A Bayesian DSGE approach,” *American economic review*, 97, 586–606.
- STOCK, J. H. AND M. W. WATSON (1999): “Forecasting inflation,” *Journal of Monetary Economics*, 44, 293–335.

Table 1: Search Query

Search Name	Search Query	Peak Month
Energy Shortages	( <i>shortages</i> N/5 <i>energy</i> ) AND <i>economics</i>	Dec-1973
Food Shortages	( <i>shortages</i> N/5 <i>food</i> ) AND <i>economics</i>	Mar-1943
Industry Shortages	( <i>shortages</i> N/5 <i>industry</i> ) AND <i>economics</i>	Aug-1942
Labor Shortages	( <i>shortages</i> N/5 <i>labor</i> ) AND <i>economics</i>	Jan-2022
Articles	<i>articles</i>	—

Topic Sets	Components
<i>shortages</i>	shortage* OR bottleneck* OR scarcit* OR rationing*
<i>energy</i>	oil OR gas OR coal OR fuel OR fuels OR gasoline OR energy OR heating OR petroleum OR electricity OR refinery OR pipeline OR petrol
<i>food</i>	food OR wheat OR meat OR milk OR crop OR crops OR grain OR farm OR agriculture OR famine OR feed OR farmer OR farmers OR water OR fertilizer OR drought
<i>industry</i>	steel OR copper OR iron OR metal* OR automotive OR textile OR machinery OR equipment OR transportation OR railway OR airline OR freight OR shipping OR transit OR deliveries OR shipment* OR ships OR chip* OR semiconductor* OR infrastructure OR materials OR distribution OR car OR cars OR parts OR goods OR material OR auto OR computer OR ‘supply chain’ OR components
<i>labor</i>	labor OR workers OR job* OR work OR employment OR manpower OR worker OR staff OR professional* OR technician* OR staffing OR skills OR workforce OR personnel OR strike* OR union*
<i>economics</i>	economic OR industr* OR production OR manufactur* OR economy OR trade OR commerce OR business OR budget OR tax OR fiscal OR corporation OR market OR price OR capacity OR company OR demand OR sales OR factory OR wages OR suppl*
<i>articles</i>	the AND be AND to AND of AND and AND at AND in

Note: The truncation character (\*) denotes a search including all possible endings of a word, e.g. “scarcit\*” includes “scarcity” and “scarcities”.

Table 2: Largest Shortage Spikes, 1900-2023

Month	Index	Surprise (st.dev.)	Event
Jan-1903	174	2.84	Nationwide coal shortages
Dec-1906	191	2.58	Shortage of coal and freight cars in Midwest
Nov-1916	235	2.54	Nationwide coal shortages
Jan-1918	578	7.63	Fuel and coal shortages
Dec-1919	300	2.75	Fuel and coal shortages due to war, strikes
Jul-1920	317	2.32	Freight car shortage affects coal and steel transportation
Aug-1922	281	3.24	Coal shortage due to strikes
Aug-1930	93	1.28	Drought leads to food and water shortages
Jul-1934	134	1.70	Strike by Teamsters unions in the West Coast
Sep-1939	152	1.81	Steel shortage due to the beginning of WW2
Aug-1941	516	3.30	War-related energy, materials and labor shortages
Jan-1943	1085	4.87	War-related oil, labor and food shortages
Jan-1945	563	4.17	War-related widespread shortages
Aug-1945	556	3.46	Labor shortages at the end of war
May-1946	589	4.31	Strikes by coal workers and fuel shortages
Jan-1948	460	4.01	Metal, fuel and food shortages
Feb-1950	226	2.30	Coal shortages amid strikes
Jan-1951	286	2.44	Labor shortages due to demand from defense industries
Jan-1952	271	2.87	Nationwide and worldwide shortages
Dec-1956	172	1.80	Oil shortages due to Suez crisis
Dec-1973	1085	11.28	Gasoline shortages due to 1973 oil crisis
Jan-1975	267	2.02	Concerns about gasoline rationing
Feb-1977	422	6.81	Carter's appeal on energy conservation
Feb-1978	158	1.59	Concerns about energy shortages
May-1979	579	4.81	Concerns about energy shortages
Aug-1981	153	1.18	Gasoline shortages due to 1979 oil crisis
Aug-1990	160	1.47	Concerns about energy shortages
Apr-2020	297	4.58	Medical shortages due to COVID-19 pandemic
Jan-2022	554	2.69	Labor shortages
Oct-2023	213	1.20	UAW strike and food shortages in Gaza

Note: The table lists the 30 largest spikes in the shortage index. For this table, the spikes are identified using the residuals of an autoregression and a condition on local maxima (see footnote 4).

Table 3: Validation of the Shortage Index

Set	Share of Total Newspaper Articles in Set	Validation Sample	Articles Mentioning Actual Shortages	Type I Error	Type II Error
Shortages AND Topic Words $\mathcal{S}$	1.58%	872	817	6.30%	–
Not Shortages $\mathcal{A} \setminus \mathcal{S}$	98.42%	298	1	–	0.33%
All Shortages $\mathcal{B}$	2.93%	334	284	14.97%	–

Note: Validation of the Shortage Index using a sample of newspaper articles used or not used to construct the index.

Table 4: Predicted Effect of Shortages on Prices and Quantities (4-quarters ahead)

	(1)		(2)		(3)	
	1950Q1–2023Q4		1950Q1–2014Q4		2015Q1–2023Q4	
	Prices	Quantities	Prices	Quantities	Prices	Quantities
GDP	0.19** (2.04)	-0.13 (-1.32)	0.22*** (3.54)	-0.21*** (-3.13)	1.02*** (4.12)	0.15 (0.57)
PCE Durables	0.25* (1.70)	-0.25*** (-3.13)	0.37*** (8.14)	-0.33*** (-3.99)	1.10 (1.31)	1.61* (1.73)
PCE Nondurables	0.31*** (2.94)	-0.15 (-1.61)	0.37*** (3.64)	-0.23*** (-2.78)	0.83* (1.85)	0.23 (0.52)
PCE Services	0.18*** (4.14)	-0.02 (-0.24)	0.19*** (4.79)	-0.05 (-0.50)	1.17*** (6.85)	0.90** (2.47)
Investment	0.28*** (2.70)	-0.25*** (-3.51)	0.32*** (5.73)	-0.31*** (-5.15)	1.27*** (3.21)	0.05 (0.08)
Obs.	292	292	260	260	32	32

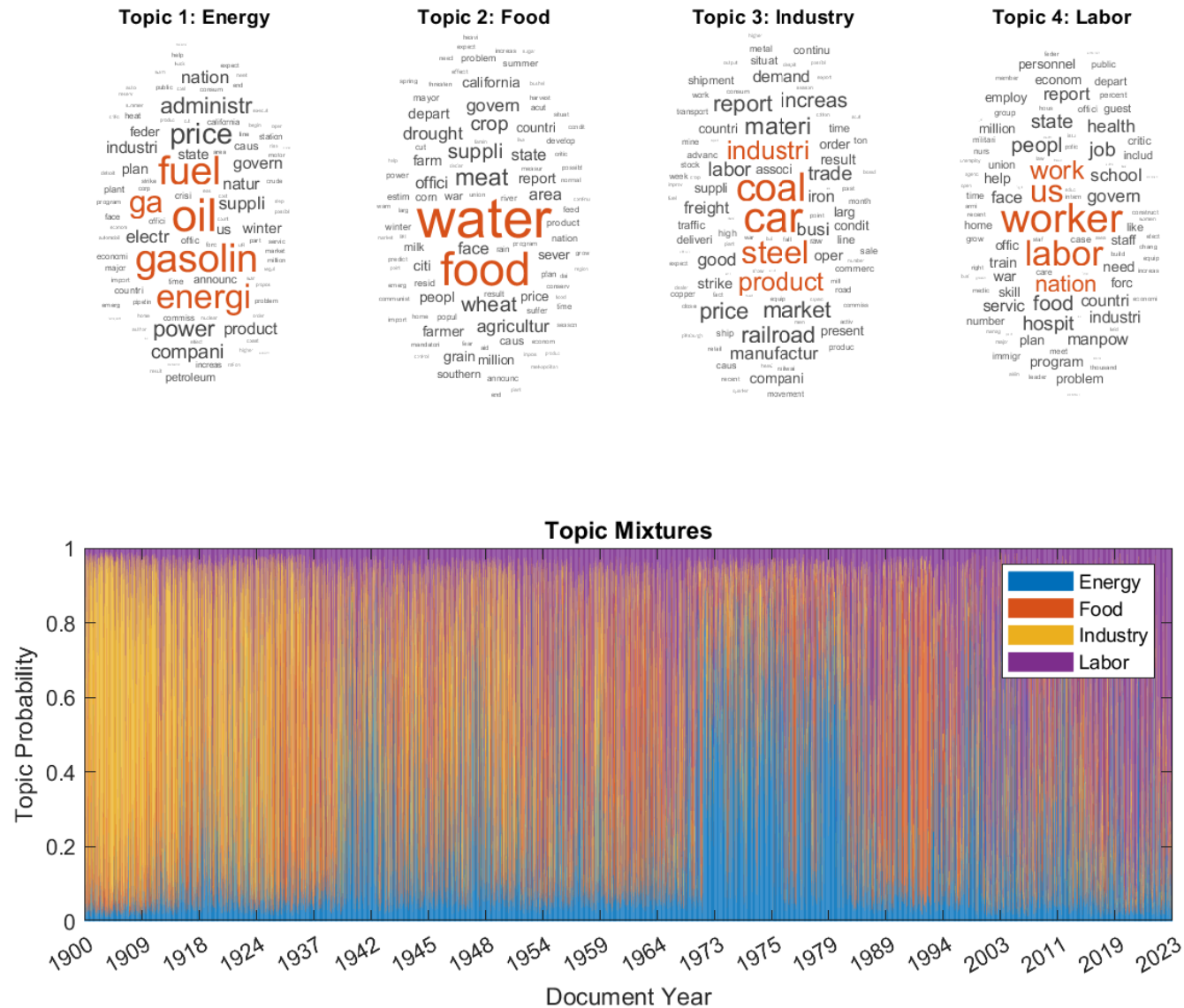
Note: The table reports standardized coefficients of predictive regressions of economic activity and inflation on the shortage index. The dependent variable for each regression is the log difference between  $t+4$  and  $t$  of the variable listed in each row, both its real quantity and its associated price deflator. Each regression includes as controls the quarterly changes of the dependent variable and the associated economic indicator or price deflator, contemporaneously and with three lags. Data are quarterly. The full sample, reported in the first two columns, runs from 1950Q1 to 2023Q4. We also partition the sample into two periods: a “pre-Covid” period that runs from 1950Q1 to 2014Q4, and a “Covid” period which runs from 2015Q1 to 2023Q4. Heteroskedasticity and autocorrelation consistent t-statistics are reported in parentheses and computed according to [Newey and West \(1987\)](#). \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table 5: Predicted Effect of Shortages on Prices and Quantities (with controls)

	(1) 1950Q1–2023Q4		(2) 1950Q1–2014Q4		(3) 2015Q1–2023Q4	
	Prices	Quantities	Prices	Quantities	Prices	Quantities
GDP	0.19** (2.04)	-0.13 (-1.32)	0.22*** (3.54)	-0.21*** (-3.13)	1.02*** (4.12)	0.15 (0.57)
Oil	0.16 (1.57)	-0.22** (-2.33)	0.22*** (5.17)	-0.35*** (-4.17)	0.92*** (3.12)	0.03 (0.14)
Commodities	0.17* (1.69)	-0.25*** (-2.67)	0.24*** (6.38)	-0.37*** (-4.39)	1.06*** (3.65)	0.13 (0.58)
Wages	0.17* (1.68)	-0.19** (-2.19)	0.24*** (7.45)	-0.29*** (-4.21)	0.29 (0.67)	-0.06 (-0.12)
Inf. Exp. (10 Yr.)	0.67*** (5.52)	0.23 (1.21)	-0.06 (-0.47)	-0.40*** (-3.54)	1.15*** (4.29)	0.14 (0.17)

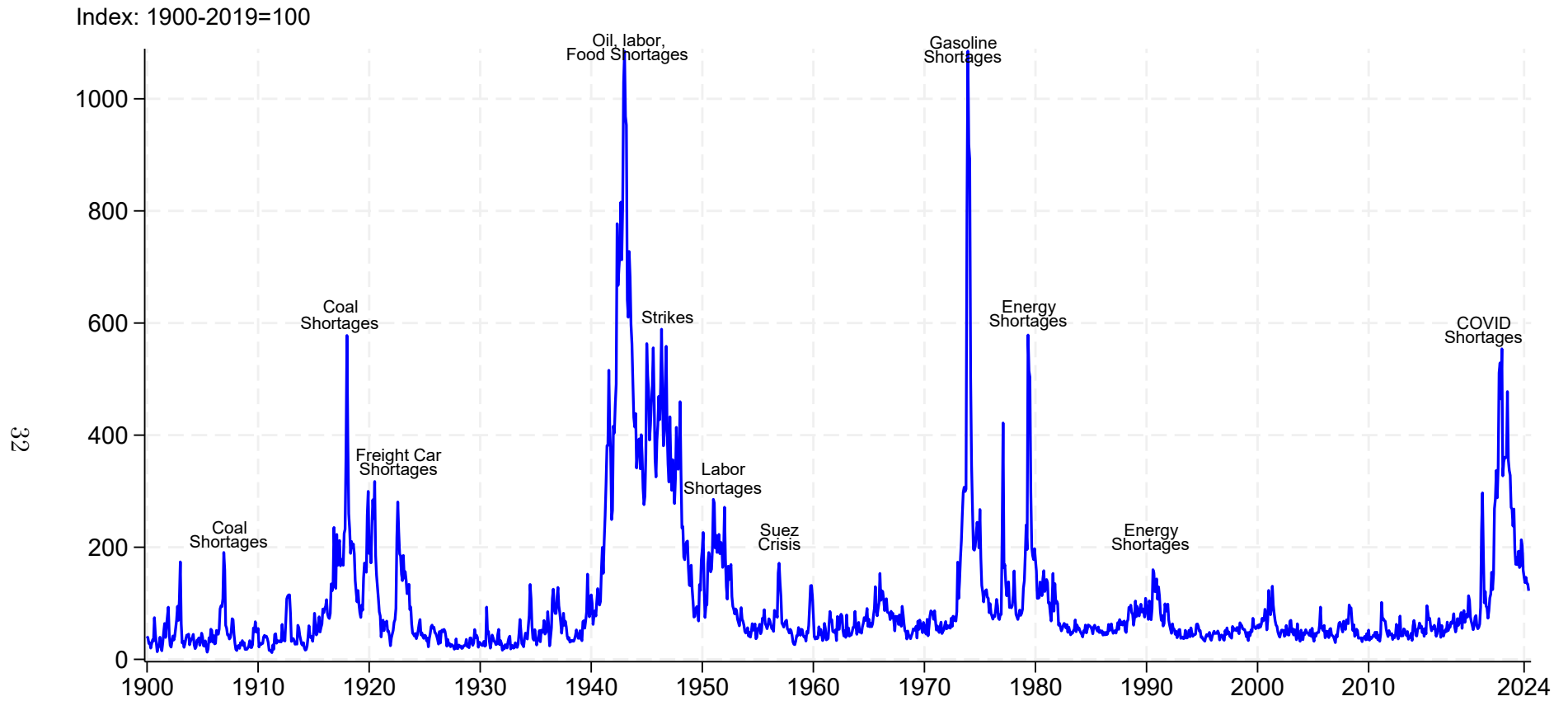
Note: The table reports standardized coefficients of predictive regressions of economic activity and inflation on the shortage index. The dependent variable for each regression is the log difference between  $t + 4$  and  $t$  of GDP, both real GDP and GDP deflator. For 10-year inflation expectations, the dependent variable for the price regressions is detrended by 10-year inflation expectations, and so we do not use this on the right-hand side. Each regression includes as controls the quarterly changes of the dependent variable and the associated economic indicator or price deflator, contemporaneously and with three lags. In each row, we add the listed variable as an additional control. Data are quarterly. The full sample, reported in the first two columns, runs from 1950Q1 to 2023Q4. We also partition the sample into two periods: a “pre-Covid” period that runs from 1950Q1 to 2014Q4, and a “Covid” period which runs from 2015Q1 to 2023Q4. Due to data limitations, the wage regressions start in 1964Q1, and the inflation expectation regressions start in 1982Q1. Heteroskedasticity and autocorrelation consistent t-statistics are reported in parentheses and computed according to [Newey and West \(1987\)](#). \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Figure 2: Topic Classification for the Index



Note: Topic classification for a large sample of articles used in the construction of the Shortage Index. The top four rows indicate the four most influential topics chosen by LDA analysis on the articles. The bottom rows show topic probability by article, sorting the articles by time.

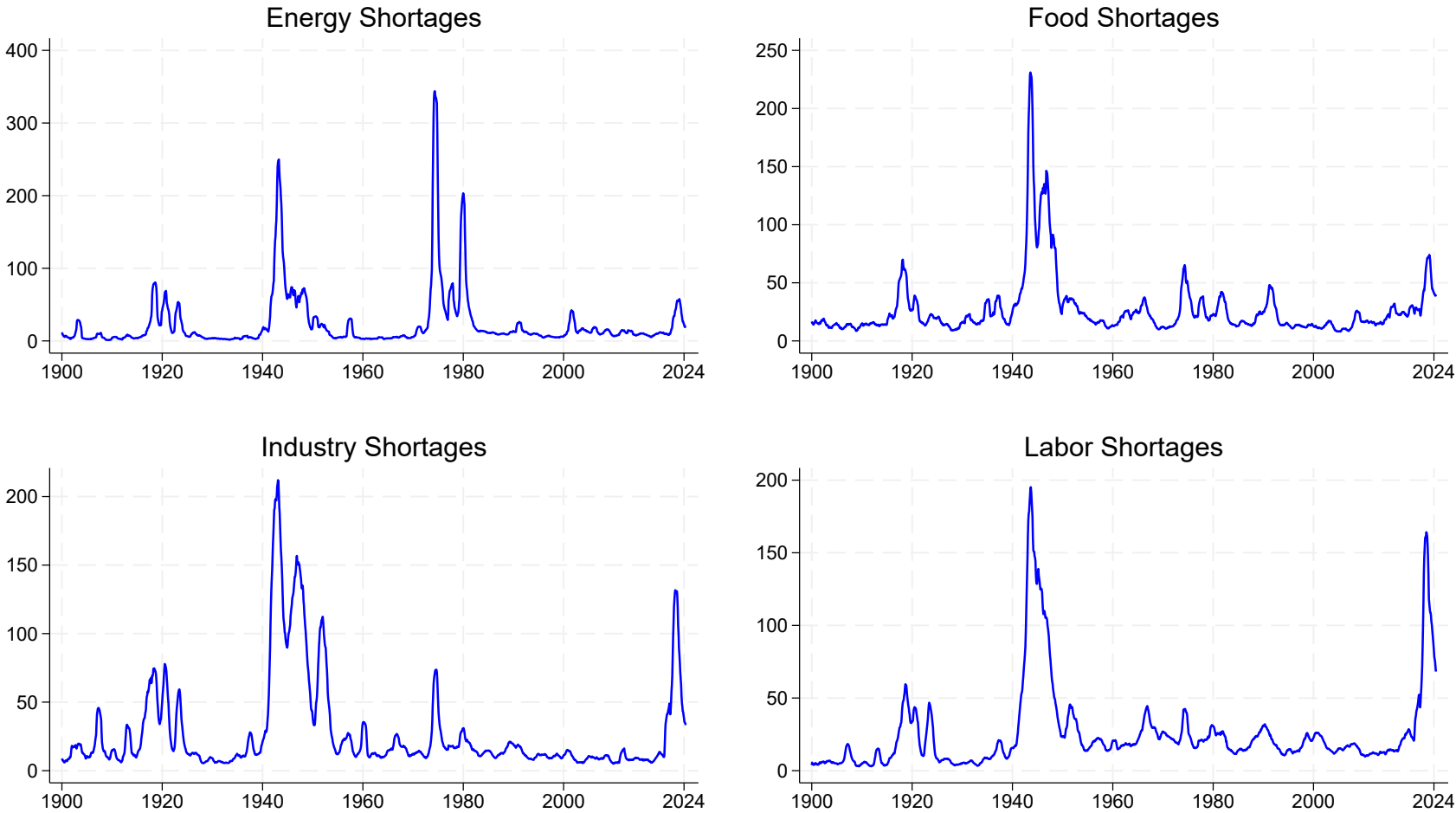
Figure 3: The Shortage Index



Note: The figure shows the shortage index from January 1900 through June 2024.



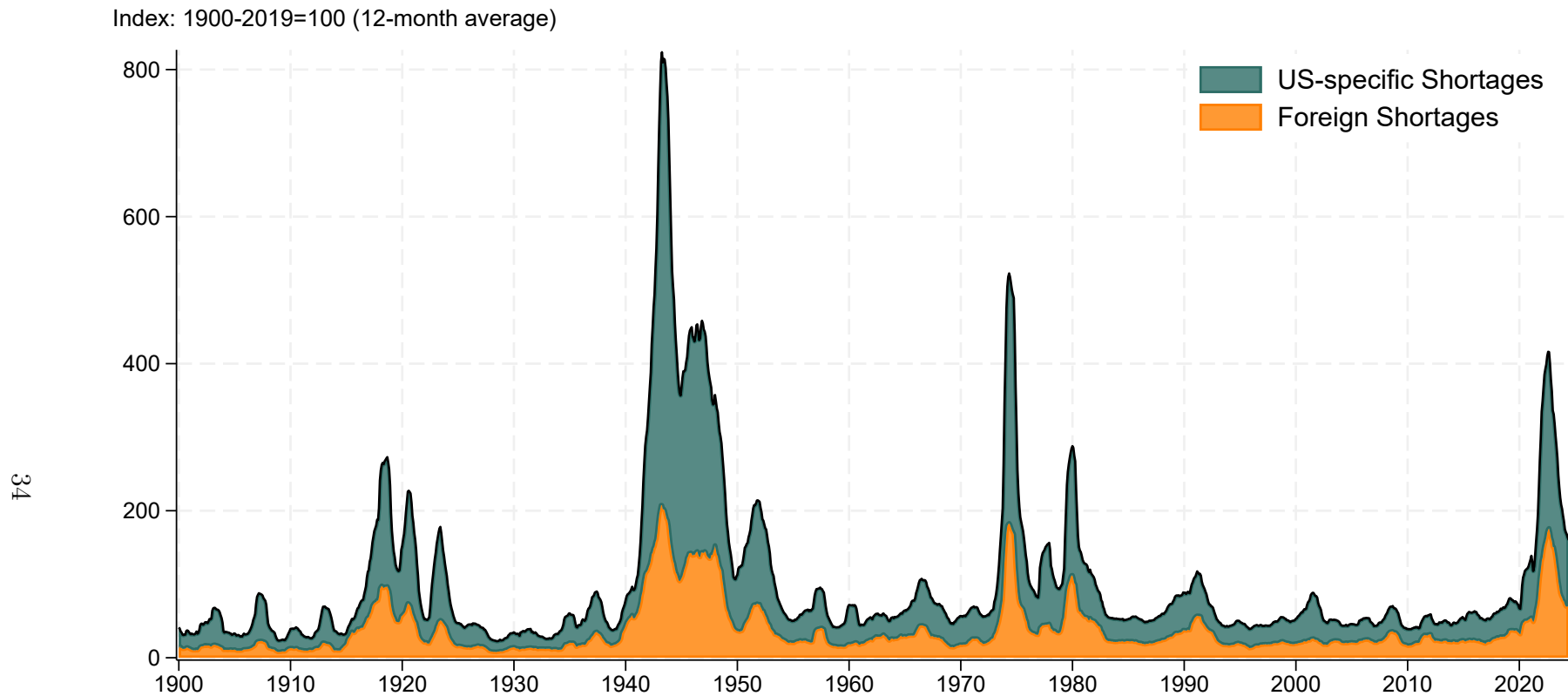
Figure 4: Decomposition of the Shortage Index by Category



12-month moving average. Scale Equal to Contribution to Total Index.

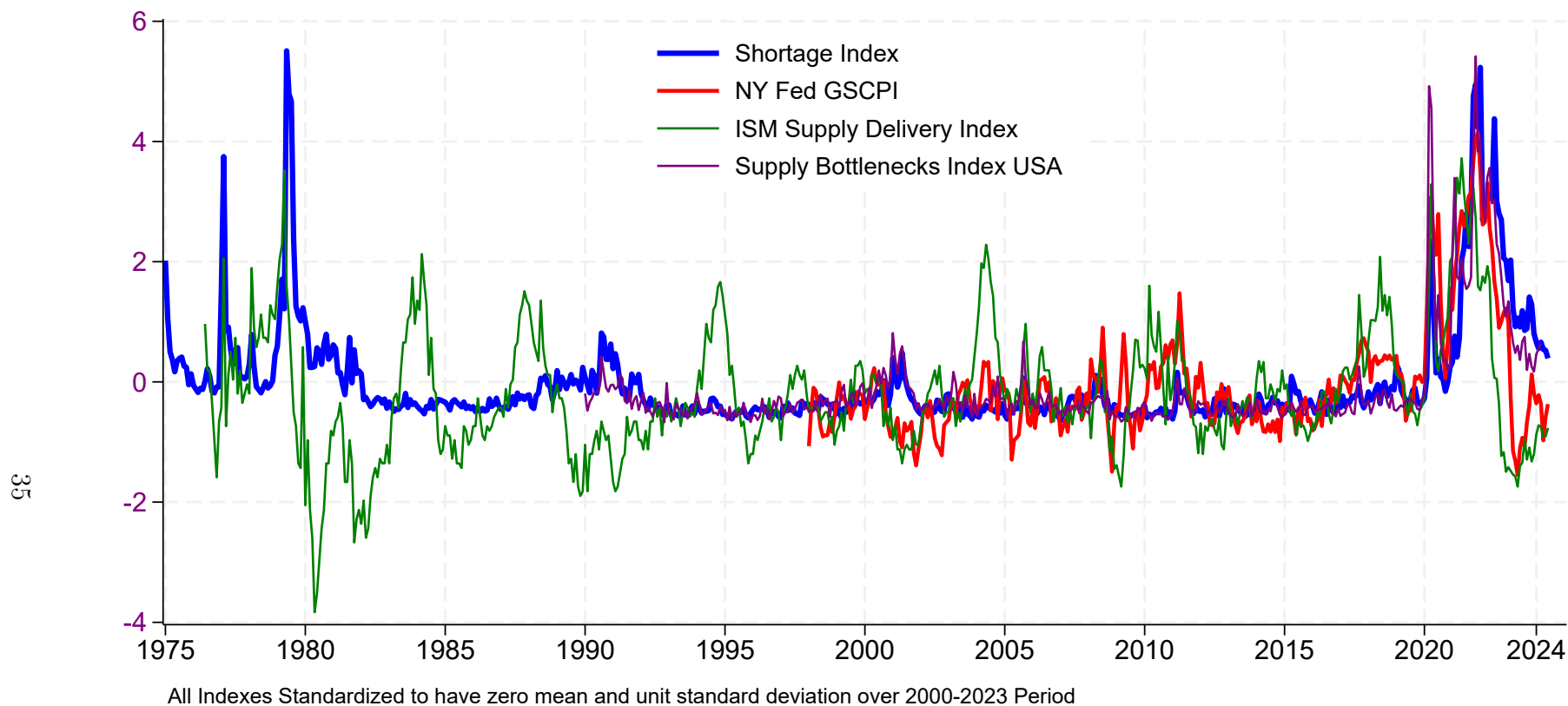
Note: The figure illustrates the four categories adding up to the total shortage index. The total index is indexed to 100 in the 1900-2019 period.

Figure 5: Shortage Index: US-specific and Foreign Components



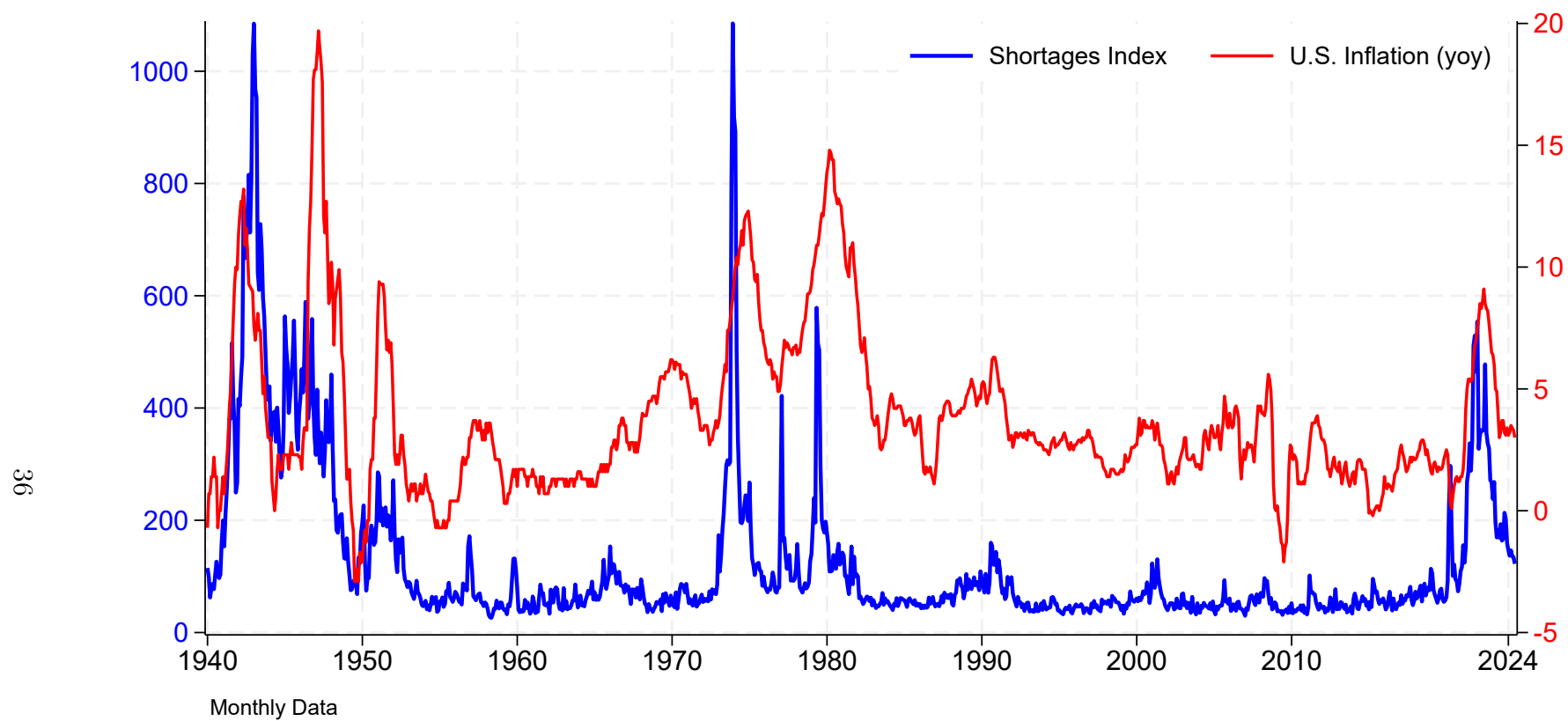
Note: The figure breaks shortages down into a US-specific and a foreign component. The foreign component is proportional to the share of newspaper articles mentioning shortages within 50 words of the name of a foreign country or city, or of the words “foreign,” “abroad,” “global,” “worldwide,” “import,” and “export.”

Figure 6: Comparison to Other Supply Constraints Measures



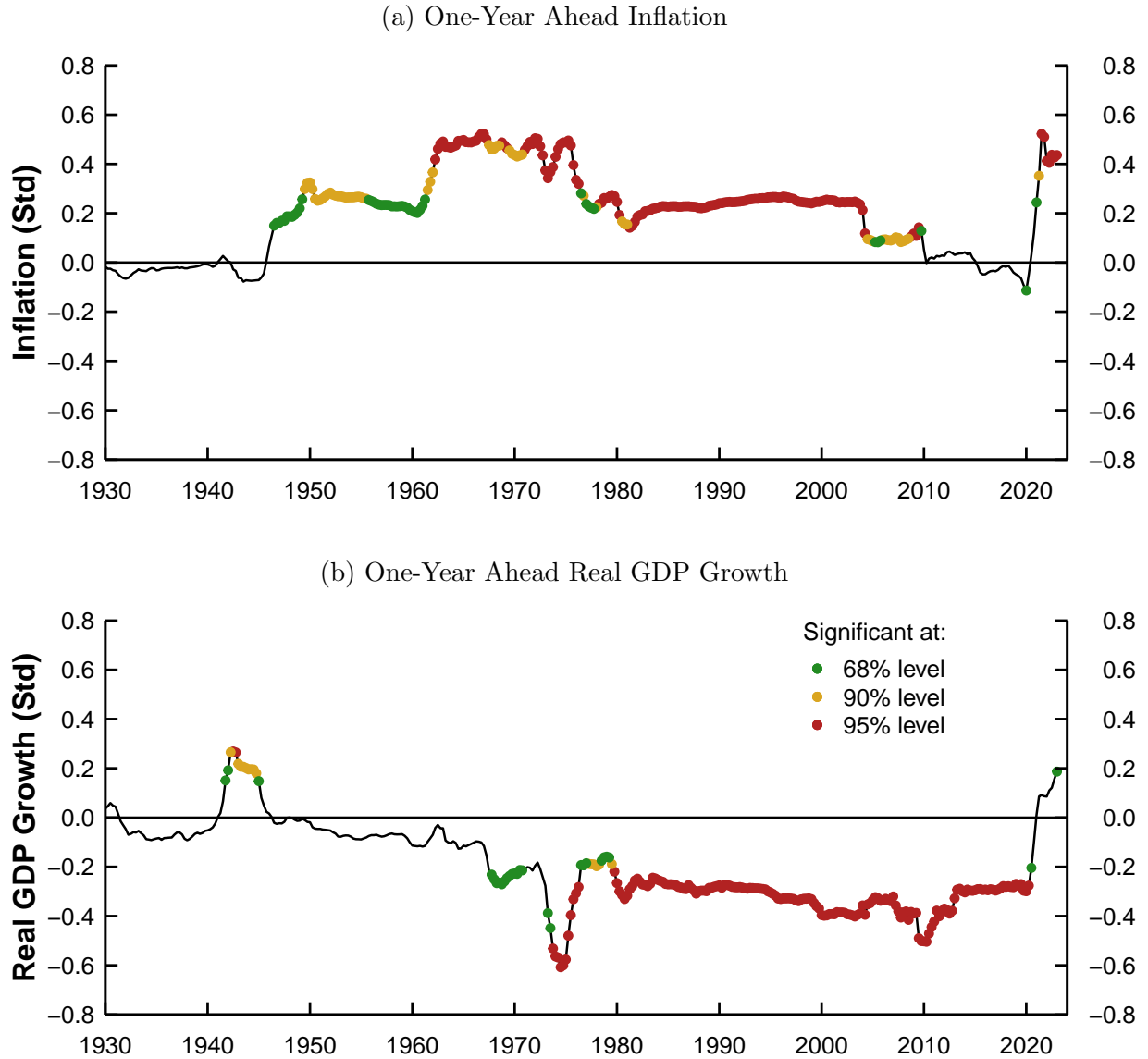
Note: The figure compares the shortage index to alternative measures of supply constraints from January 1975 through June 2024. The ISM Supply Delivery Index is computed as the share of respondents reporting longer delivery times plus half the share of respondents reporting no change in delivery times. The Supply Bottlenecks Index is the U.S. Supply Bottlenecks Index from [Burriel et al. \(2023\)](#).

Figure 7: Shortages and U.S. Inflation



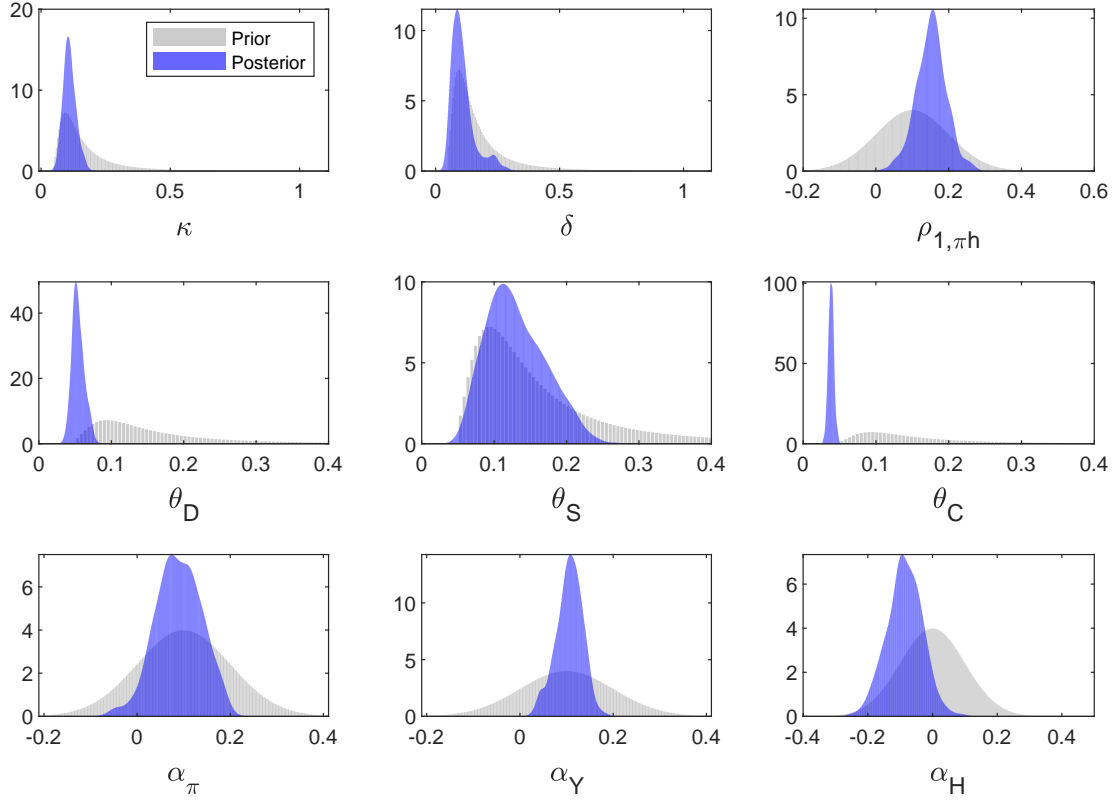
Note: The figures compares the shortage index (left) with U.S. inflation (right), from January 1940 through June 2024.

Figure 8: Effect of Shortages on Inflation and Real GDP Growth (30-Year Window)



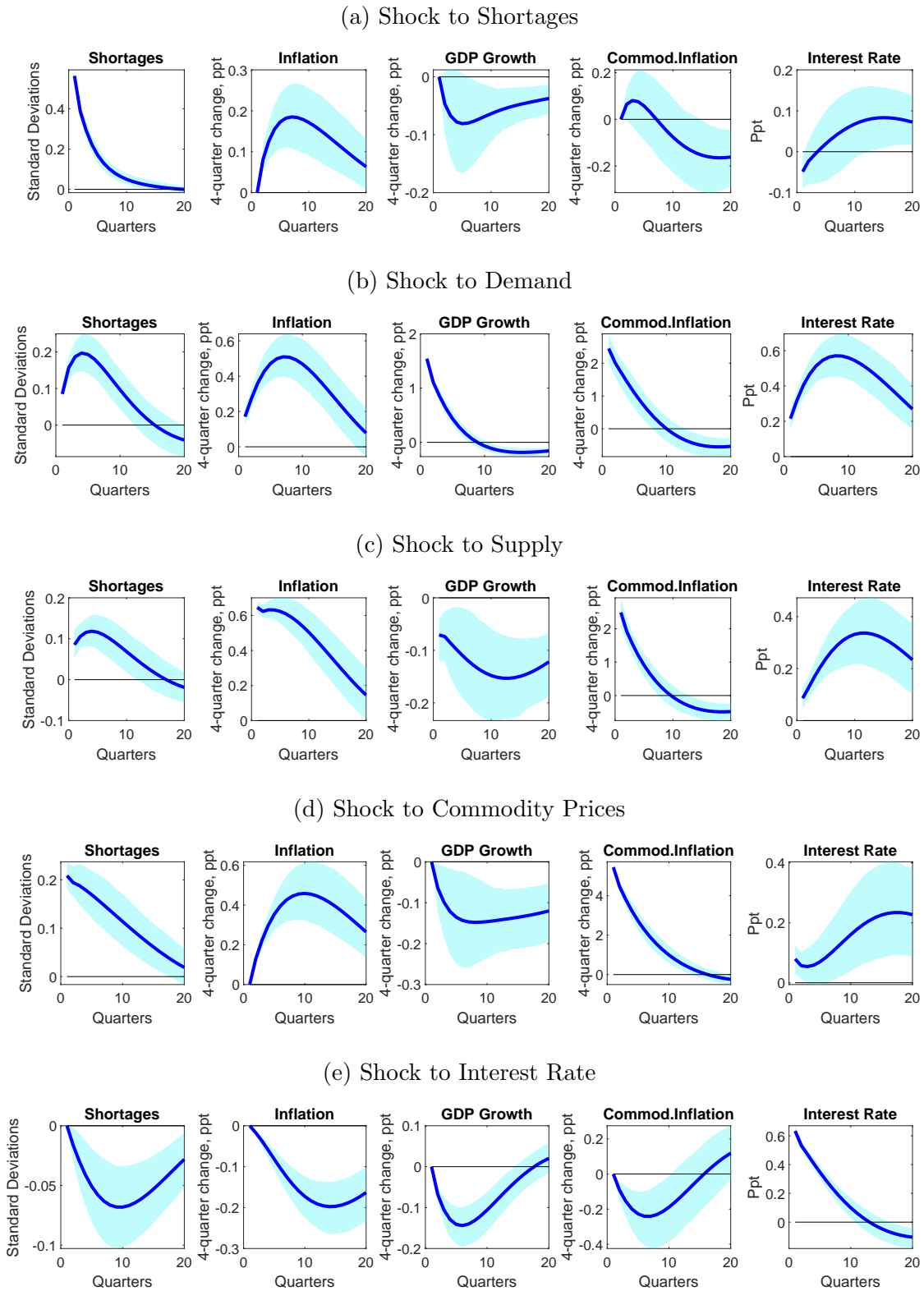
Note: The solid black line shows the time-varying estimated effect of shortages on inflation (top panel) and real GDP growth (bottom panel). The estimates are based on regressions using rolling 30-year windows. In each regression, the dependent variable is the 4-quarter ahead difference in log real GDP per capita or the 4-quarter ahead difference in log GDP deflator. On the right-hand side, the main explanatory variable is our shortage index. As controls, we include the one-quarter change in both log real GDP per capita and log GDP deflator, in quarter  $t$  plus three lags, and allow for up to three quarters of autocorrelation.

Figure 9: Prior and Posterior Densities of Structural VAR Parameters



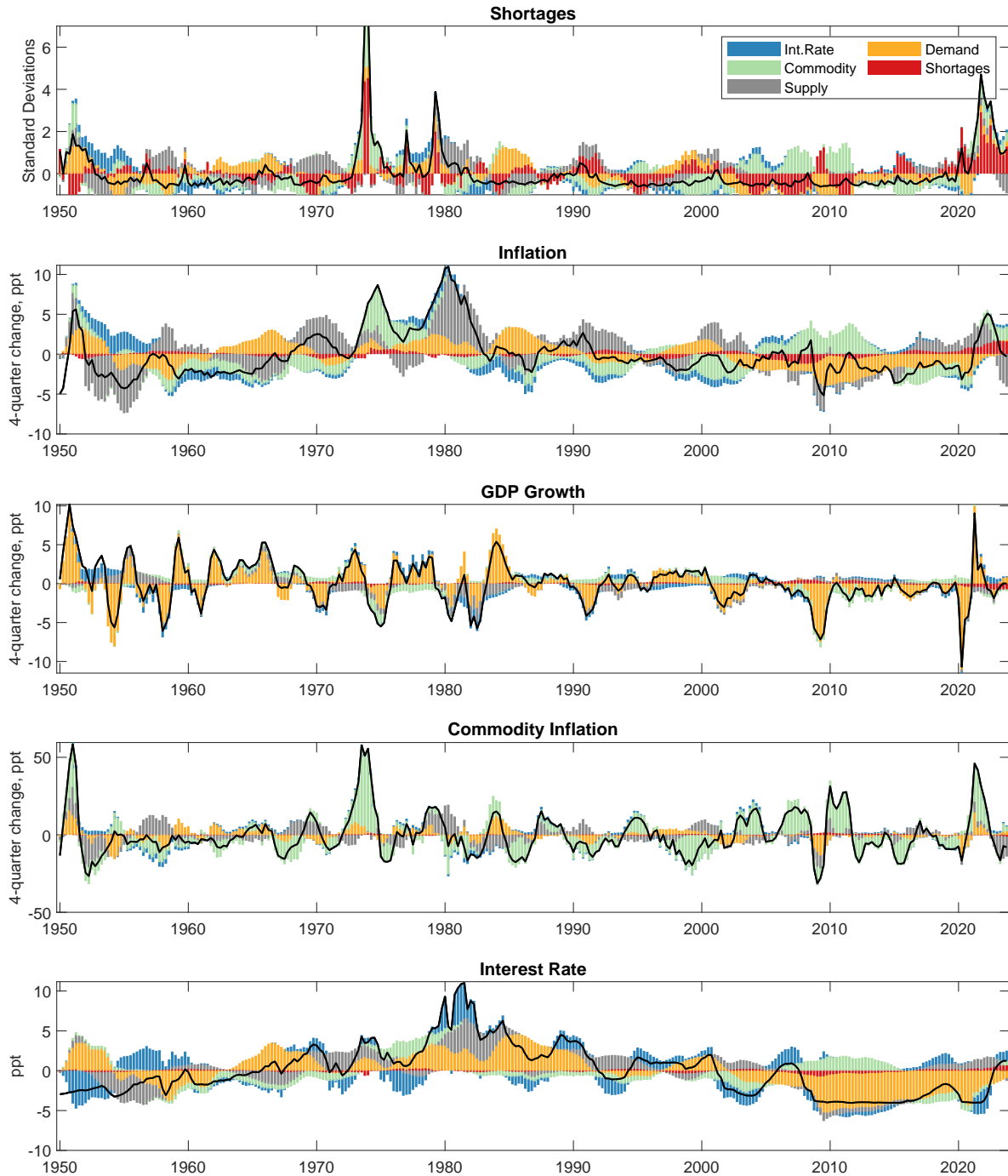
Note: Prior and posterior densities of selected parameters of the structural VAR model.

Figure 10: Effects of Shortages on US Activity and Inflation in the Estimated VAR Model



Note: Impulse Response in VAR to Estimated Shocks. All shocks are one-standard deviation in size. Solid lines denote the response at the posterior mean. Shaded areas denote 80 percent confidence intervals.

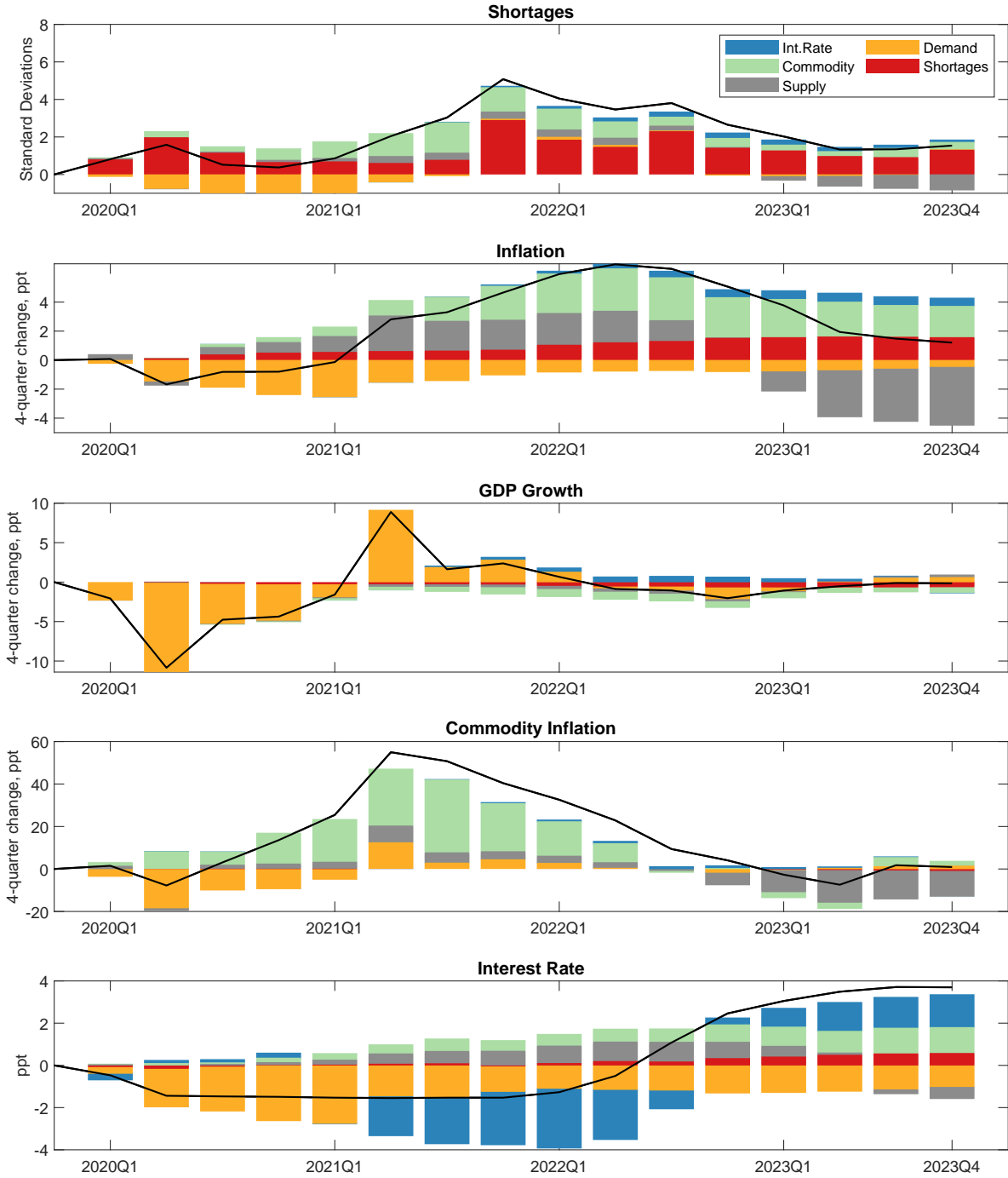
Figure 11: Historical Decomposition of Estimated VAR Model: Full Sample



Note: The figure decomposes movements in shortages, economic activity and inflation in the contribution of shortage shocks, demand shocks, supply shocks, and commodity price shocks. All variables are expressed in deviation from their sample mean.

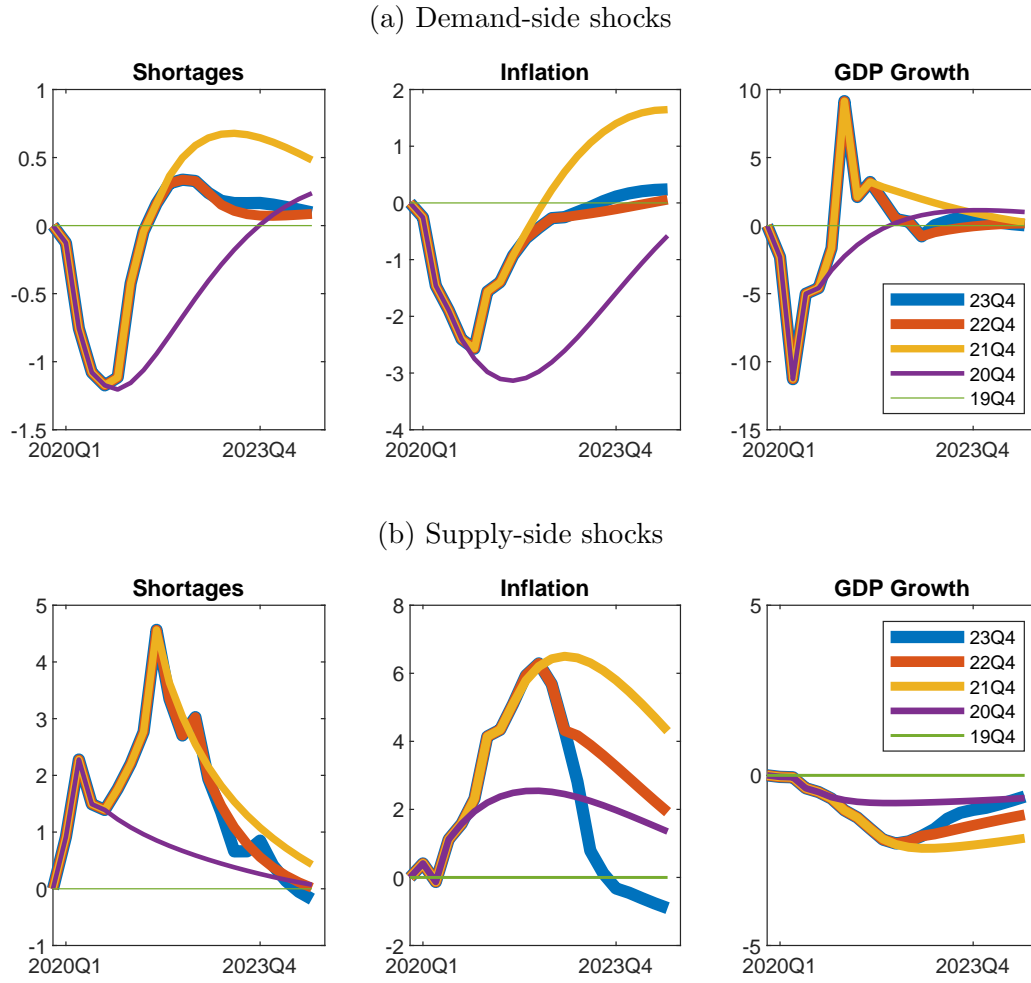


Figure 12: Historical Decomposition of Estimated VAR Model: 2020-2023



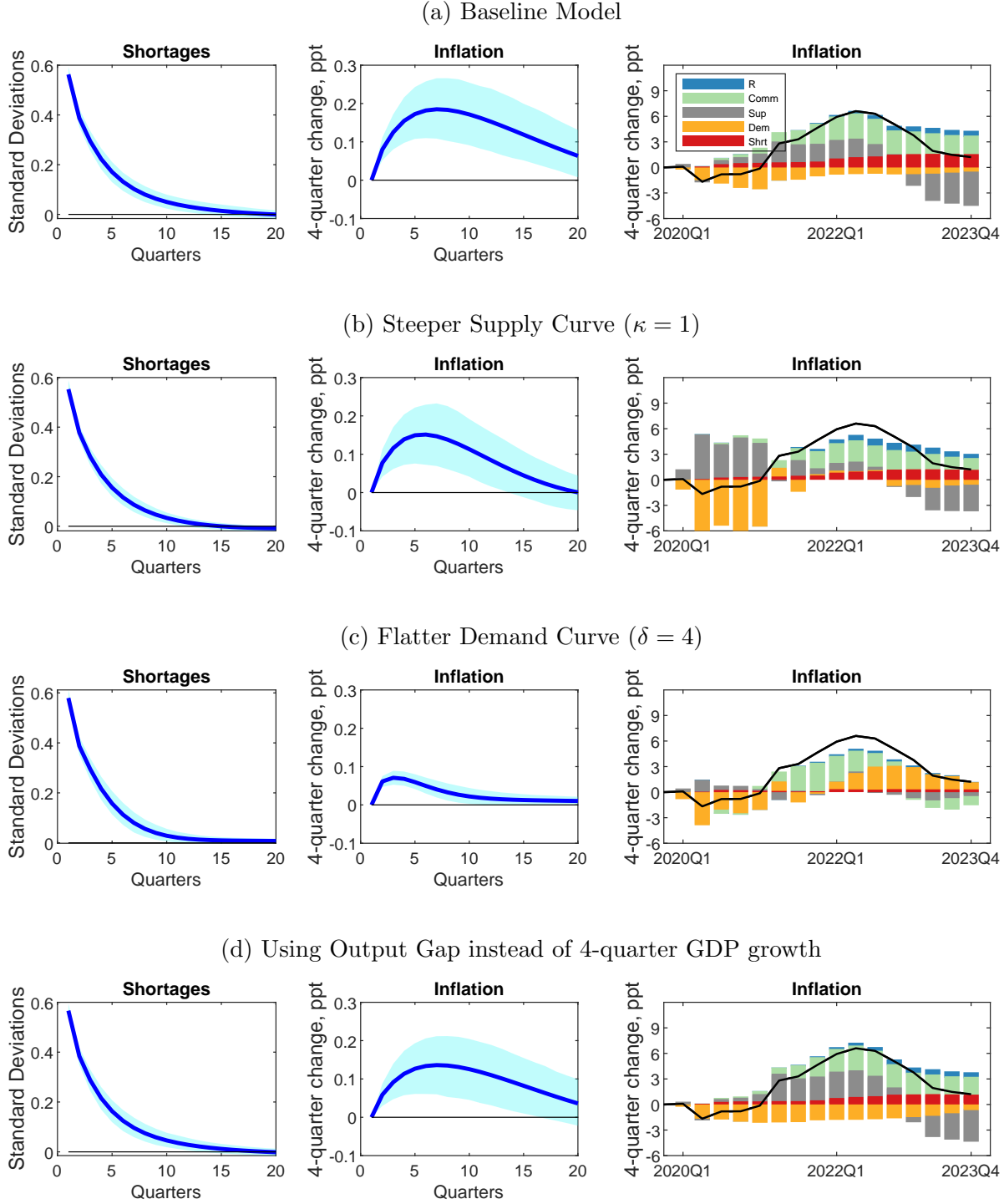
Note: The figure decomposes movements in shortages, economic activity and inflation since 2020 in the contribution of shortage shocks, demand shocks, supply shocks, and commodity price shocks. All variables are expressed in deviation from their 2019Q4 value.

Figure 13: Contribution to Inflation of Demand-side and Supply-side Forces: 2020-2024



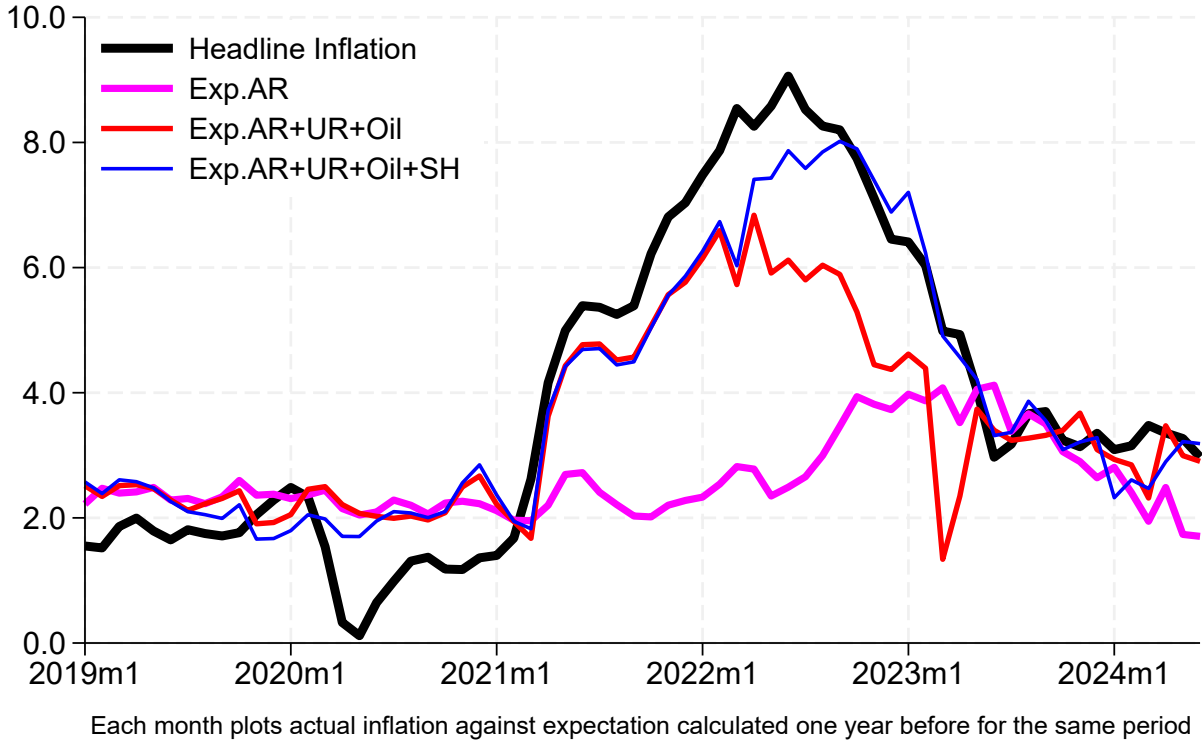
Note: The figure shows predicted contribution between 2019:Q4 and 2024:Q4 of demand-side forces (first row) and supply-side forces (second row) to inflation and GDP growth. Demand forces are the sum of demand and monetary shocks. Supply forces are the sum of shortage, commodity, and supply shocks. Variables are plotted in deviation from their 2019:Q4 value. The model is estimated through 2023:Q4. Last period is 2024:Q4.

Figure 14: Comparison of baseline VAR model with alternative versions



Note: Models (b) and (c) are estimating after fixing  $\kappa = 1$  and  $\delta = 4$ , respectively. Model (d) replaces 4-quarter GDP growth with the output gap constructed using the CBO measure of the output GDP.

Figure 15: Inflation and its Forecasts



Note: The figure plots realizations of CPI inflation with corresponding forecasts constructed 12 months earlier. Following [Stock and Watson \(1999\)](#), the series are aligned so that the vertical distance between the plot of inflation and the forecast represents the forecast error. The model Exp.AR predicts one-year ahead inflation using current and 12 lags of inflation. The model Exp.AR+UR+Oil includes unemployment and the 12-month change in oil prices. The model Exp.AR+UR+Oil+SH adds shortages to the previous model.

# Appendix

## A Appendix Tables

Table A.1: Predicted Effect of Shortages on Prices and Quantities (1-quarter ahead)

	(1) 1950Q1–2023Q4		(2) 1950Q1–2014Q4		(3) 2015Q1–2023Q4	
	Prices	Quantities	Prices	Quantities	Prices	Quantities
GDP	0.10** (2.36)	-0.01 (-0.18)	0.11*** (3.22)	-0.06 (-1.09)	0.91*** (5.60)	-0.10 (-0.45)
PCE Durables	0.12 (1.58)	-0.06 (-1.13)	0.19*** (4.32)	-0.12** (-2.13)	0.07 (0.15)	1.21 (1.28)
PCE Nondurables	0.28*** (3.67)	-0.06 (-0.71)	0.31*** (3.83)	-0.09 (-1.13)	0.52* (1.85)	0.27 (0.80)
PCE Services	0.13*** (4.35)	-0.02 (-0.33)	0.14*** (4.70)	-0.06 (-0.91)	0.71*** (4.77)	0.57** (2.14)
Investment	0.17*** (3.79)	-0.02 (-0.38)	0.18*** (5.13)	-0.05 (-1.14)	0.89*** (5.04)	-0.03 (-0.07)
Obs.	295	295	260	260	35	35

Note: The table reports standardized coefficients of predictive regressions of economic activity and inflation on the shortage index. The dependent variable for each regression is the log difference between  $t + 1$  and  $t$  of the variable listed in each row and its associated price deflator. Each regression includes as controls the quarterly changes of the dependent variable and the associated economic indicator or price deflator, contemporaneously and with three lags. Data are quarterly. The full sample, reported in the first two columns, runs from 1950Q1 to 2023Q4. We also partition the sample into two periods: a “pre-Covid” period that runs from 1950Q1 to 2014Q4, and a “Covid” period which runs from 2015Q1 to 2023Q4. Heteroskedasticity and autocorrelation consistent t-statistics are reported in parentheses and computed according to [Newey and West \(1987\)](#). \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table A.2: Predicted Effect of Shortages on Prices and Quantities (8-quarters ahead)

	(1) 1950Q1–2023Q4		(2) 1950Q1–2014Q4		(3) 2015Q1–2023Q4	
	Prices	Quantities	Prices	Quantities	Prices	Quantities
GDP	0.17** (2.53)	-0.08 (-1.14)	0.15** (2.54)	-0.12* (-1.72)	0.46 (0.96)	0.96 (1.37)
PCE Durables	0.28*** (3.06)	-0.16** (-1.98)	0.32*** (6.70)	-0.21* (-1.94)	1.41 (1.31)	-0.29 (-0.30)
PCE Nondurables	0.25*** (2.74)	-0.02 (-0.27)	0.25*** (2.84)	-0.08 (-1.04)	1.29*** (3.24)	0.12 (0.36)
PCE Services	0.19*** (3.52)	-0.00 (-0.05)	0.18*** (3.83)	0.01 (0.10)	1.13*** (5.28)	1.14** (2.68)
Investment	0.26*** (4.42)	-0.19*** (-2.96)	0.25*** (4.71)	-0.22*** (-3.63)	0.76 (1.52)	0.44 (0.72)
Obs.	288	288	260	260	28	28

Note: The table reports standardized coefficients of predictive regressions of economic activity and inflation on the shortage index. The dependent variable for each regression is the log difference between  $t+4$  and  $t$  of the variable listed in each row and its associated price deflator. Each regression includes as controls the quarterly changes of the dependent variable and the associated economic indicator or price deflator, contemporaneously and with three lags. Data are quarterly. The full sample, reported in the first two columns, runs from 1950Q1 to 2023Q4. We also partition the sample into two periods: a “pre-Covid” period that runs from 1950Q1 to 2014Q4, and a “Covid” period which runs from 2015Q1 to 2023Q4. Heteroskedasticity and autocorrelation consistent t-statistics are reported in parentheses and computed according to [Newey and West \(1987\)](#). \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table A.3: VAR Model: Estimated Parameters

Parameter	Type	Prior		Posterior			
		Mean	St. Dev	Mean	St. Dev	10%	90%
$\sigma_{\varepsilon_d}$	Inv. Gamma	<b>1.00</b>	0.50	<b>1.56</b>	0.06	1.49	1.65
$\sigma_{\varepsilon_s}$	Inv. Gamma	<b>1.00</b>	0.50	<b>0.65</b>	0.02	0.62	0.68
$\sigma_{\varepsilon_c}$	Inv. Gamma	<b>1.00</b>	0.50	<b>5.44</b>	0.21	5.17	5.72
$\sigma_{\varepsilon_h}$	Inv. Gamma	<b>1.00</b>	0.50	<b>0.56</b>	0.02	0.53	0.59
$\sigma_{\varepsilon_r}$	Inv. Gamma	<b>1.00</b>	0.50	<b>0.63</b>	0.03	0.60	0.67
$\rho_{1,HH}$	Beta	<b>0.65</b>	0.10	<b>0.69</b>	0.04	0.63	0.73
$\rho_{1,YY}$	Beta	<b>0.65</b>	0.10	<b>0.75</b>	0.04	0.70	0.80
$\rho_{1,PP}$	Beta	<b>0.65</b>	0.10	<b>0.88</b>	0.03	0.84	0.92
$\rho_{1,CC}$	Beta	<b>0.65</b>	0.10	<b>0.82</b>	0.03	0.78	0.86
$\rho_{1,RR}$	Beta	<b>0.65</b>	0.10	<b>0.85</b>	0.03	0.82	0.89
$\rho_{1,PR}$	Beta	<b>0.10</b>	0.05	<b>0.02</b>	0.01	0.01	0.03
$\rho_{1,YR}$	Beta	<b>0.50</b>	0.10	<b>0.11</b>	0.03	0.08	0.14
$\delta$	Inv. Gamma	<b>0.20</b>	1.00	<b>0.11</b>	0.05	0.06	0.17
$\kappa$	Inv. Gamma	<b>0.20</b>	1.00	<b>0.11</b>	0.02	0.08	0.14
$\theta_C$	Inv. Gamma	<b>0.20</b>	1.00	<b>0.04</b>	0.00	0.03	0.04
$\theta_D$	Inv. Gamma	<b>0.20</b>	1.00	<b>0.05</b>	0.01	0.04	0.07
$\theta_S$	Inv. Gamma	<b>0.20</b>	1.00	<b>0.13</b>	0.04	0.08	0.19
$\alpha_C$	Normal	<b>0.00</b>	0.10	<b>0.02</b>	0.01	0.01	0.03
$\alpha_Y$	Normal	<b>0.10</b>	0.10	<b>0.11</b>	0.03	0.07	0.14
$\alpha_H$	Normal	<b>0.00</b>	0.10	<b>-0.09</b>	0.06	-0.16	-0.02
$\alpha_P$	Normal	<b>0.10</b>	0.10	<b>0.09</b>	0.05	0.03	0.15
$\chi_D$	Inv. Gamma	<b>2.00</b>	5.00	<b>1.57</b>	0.21	1.31	1.85
$\chi_S$	Inv. Gamma	<b>2.00</b>	5.00	<b>3.80</b>	0.50	3.15	4.46
$\rho_{1,HY}$	Normal	<b>0.00</b>	0.10	<b>0.05</b>	0.02	0.03	0.07
$\rho_{1,HP}$	Normal	<b>0.00</b>	0.10	<b>0.04</b>	0.02	0.02	0.07
$\rho_{1,HC}$	Normal	<b>0.00</b>	0.10	<b>0.01</b>	0.00	0.01	0.01
$\rho_{1,YH}$	Normal	<b>0.00</b>	0.10	<b>-0.07</b>	0.08	-0.17	0.03
$\rho_{1,YP}$	Normal	<b>0.00</b>	0.10	<b>0.11</b>	0.05	0.05	0.18
$\rho_{1,YC}$	Normal	<b>0.00</b>	0.10	<b>-0.00</b>	0.01	-0.02	0.01
$\rho_{1,PH}$	Normal	<b>0.10</b>	0.10	<b>0.15</b>	0.04	0.10	0.21
$\rho_{1,PY}$	Normal	<b>0.00</b>	0.10	<b>-0.04</b>	0.03	-0.07	-0.00
$\rho_{1,PC}$	Normal	<b>0.10</b>	0.10	<b>0.02</b>	0.00	0.01	0.02
$\rho_{1,CH}$	Normal	<b>0.10</b>	0.10	<b>0.10</b>	0.09	-0.02	0.23
$\rho_{1,CY}$	Normal	<b>0.00</b>	0.10	<b>0.04</b>	0.08	-0.07	0.14
$\rho_{1,CP}$	Normal	<b>0.00</b>	0.10	<b>-0.17</b>	0.09	-0.29	-0.05
$\rho_{1,RP}$	Normal	<b>0.00</b>	0.10	<b>0.03</b>	0.05	-0.04	0.09
$\rho_{1,RY}$	Normal	<b>0.00</b>	0.10	<b>0.00</b>	0.03	-0.03	0.04
$\rho_{1,RC}$	Normal	<b>0.00</b>	0.10	<b>-0.02</b>	0.01	-0.03	-0.01
$\rho_{1,RH}$	Normal	<b>0.00</b>	0.10	<b>0.09</b>	0.06	0.01	0.16
$\rho_{1,CR}$	Beta	<b>0.10</b>	0.05	<b>0.15</b>	0.06	0.07	0.23
$\rho_{1,HR}$	Normal	<b>0.00</b>	0.10	<b>0.03</b>	0.02	0.00	0.05
$\rho_{2,HH}$	Beta	<b>0.10</b>	0.05	<b>0.04</b>	0.02	0.02	0.07
$\rho_{2,YY}$	Beta	<b>0.10</b>	0.05	<b>0.05</b>	0.02	0.02	0.08
$\rho_{2,PP}$	Beta	<b>0.10</b>	0.05	<b>0.06</b>	0.03	0.03	0.10
$\rho_{2,CC}$	Beta	<b>0.10</b>	0.05	<b>0.03</b>	0.02	0.01	0.05
$\rho_{2,RR}$	Beta	<b>0.10</b>	0.05	<b>0.05</b>	0.02	0.02	0.08

Note: The table shows estimated parameters of the model described in Section 5. The coefficients describing lagged response of  $\pi$ ,  $y$ ,  $c$ , and  $h$  to interest rates are written with a negative sign in the model specification of the VAR model to impose a negative lagged impact of interest rates on each of these variables.

## B Additional Details of Audit Conducted with Claude AI

We used Claude AI assistant to help us with auditing the index. Claude was instructed to return a table of results, coding articles as 1 (shortage mentioned), 0 (shortage not mentioned), or 99 (unsure whether the existence of shortages was mentioned). In addition to the classification, Claude was asked to provide a brief explanation for each snippet’s coding.

Before initiating the classification, we provided Claude with examples of how we would code the snippets and ensured that the training sample included false positives, such as mentions of the lack or end of shortages. The specific prompt was as follows:

“I give you 872 snippets of text, each about 110 characters long. For each of them, can you tell me whether they mention current or prospective shortages, rationing, scarcity, or bottlenecks related to goods, labor, materials, food, or water? Just return a table with yes=1, no=0, unsure=99, and a brief explanation. For instance:

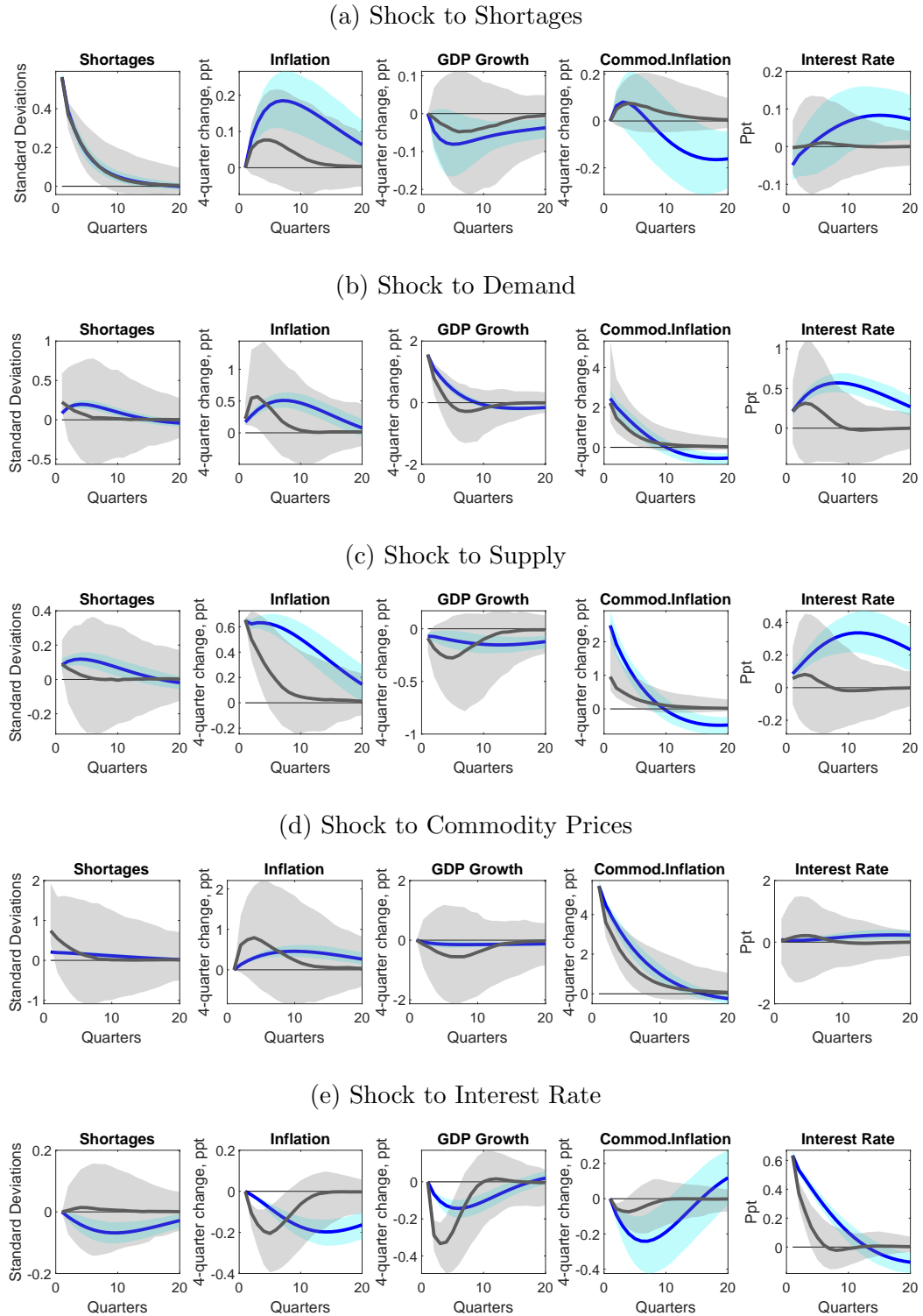
- Article 1 mentions prospective shortages since it states that steel shortages will prevail in the near future, so it is coded as 1.
- Article 2 says steel shortages caused a plant closure, so it is coded as 1.
- Article 3 says the shortage of cars is crimping coal production, so it is coded as 1.
- Article 4 mentions the shortage of cars, so it is coded as 1.
- Article 10 mentions a shortage of workspace, so it is not related to a work shortage and is coded as 0.
- Article 329 says no shortage of cars has been experienced, so it is coded as 0.”

Although using AI for validation is not foolproof, we found that Claude performed comparably to a human classifier. For example, Claude demonstrated an ability to extrapolate the context of a particular sentence to a country or person. In one instance, the snippet “Economy may be slowing, but Lowe is banking on labor shortages gradually leading to an increase...” was classified as 1 by Claude, with the explanation: “Reserve Bank [of Australia] expecting labor shortages to lead to wage growth.”

The results of the audit are in Table 3. Out of 872 articles belonging to the set  $\mathcal{S}$ , only 6.3 percent were deemed by ClaudeAI as false positives. Claude classified the snippet “...says that while there is not necessarily a shortage of people wanting to work in management” as 0 with the explanation “No shortage of people wanting to work in management.” Similarly, the snippet “a motive for mr. newt gingrich’s knife job, had no shortage of conspiracy theories, most leading to the...” was classified as 0 with the explanation “Speaker’s ouster sparked many conspiracy theories but not actual shortages.” In some cases, Claude classified as 0 articles that we would have probably classified as 1. For instance, the snippet “canada’s action today in temporarily suspending meat rationing” was classified as 0 since Claude gave more weight to the temporary suspension of the rationing rather than its existence.



Figure A.1: Effects of Shortages in the Estimated VAR Model: Prior vs Posterior



Note: Red lines and areas, prior; Blue lines and areas, posterior. Impulse Response in VAR to Estimated Shocks, Prior vs. Posterior. Posterior shocks are one-standard deviation in size. Solid lines denote the response at the mean. Shaded areas denote 80 percent confidence intervals. Impact of the shocks at the prior mean has been normalized to match the impact effect of the estimated shocks at the posterior mean.