

# Measuring Shortages Since 1900\*

Dario Caldara      Matteo Iacoviello      David Yu

April 2024

## PRELIMINARY DRAFT

### **Abstract**

This paper introduces a monthly shortage index for the United States from 1900 to 2023, constructed using a sample of approximately 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 war. We explore the relationship between the shortage index and key economic indicators, discussing potential applications for researchers, policymakers, and businesses.

**KEYWORDS:** Shortages; Textual Analysis.

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

---

\*Corresponding author: Matteo Iacoviello ([matteo.iacoviello@frb.gov](mailto:matteo.iacoviello@frb.gov)). All authors are affiliated with the Federal Reserve Board. 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 the economics discipline is the study of the scarce allocation of 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.

This paper aims to fill this gap by constructing 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. It is constructed using a sample of about 20,000 news articles per month from six major U.S. newspapers. In total, approximately 25 million articles are used in the construction of the index.

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. We hope that the resulting shortage index can provide a valuable tool for researchers, policymakers, and businesses to study the dynamics of shortages and their relationship to various economic, political, and social factors. By examining the index 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. Example include [Bernanke and Blanchard \(2023\)](#), [Chen and Houle \(2023\)](#), [Lamont \(1997\)](#), [Pitschner \(2022\)](#), [Benigno et al. \(2022\)](#), [Burriel et al. \(2023\)](#).

The novelty of our approach is that our index is the first spanning over 125 years.

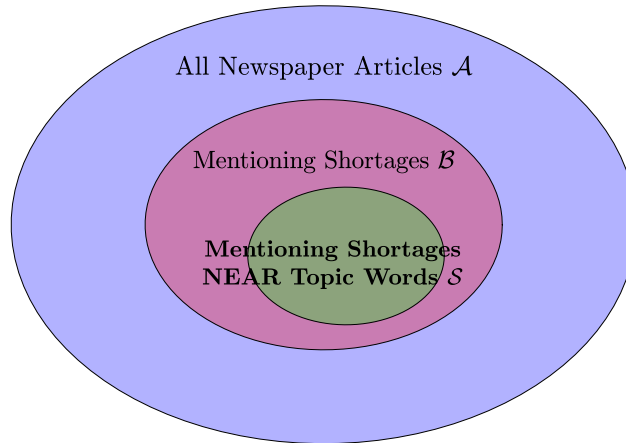
The remainder of the paper is structured as follows. Section 2 discusses the construction of the index, presents the index and discusses its evolution over time. Section 3 validates the index. Section 4 explores the relationship between the shortage index and inflation in the United States. Section 5 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 discuss 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 of the intensity of shortages of raw materials, goods, services and labor in the United States. It is constructed using a sample of about 20,000 news articles per month from 1900 through today—for a total of about 25 million articles—published in 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 discussing energy, food, industry or labor shortages (the set  $\mathcal{S}$ ), normalized by the total number of articles (the set  $\mathcal{A}$ ). Accordingly, higher values of the index indicates higher intensity of shortages.



Grouping of newspaper articles for the construction of the shortages index.

To build the search query for the index, we start by analyzing the text of a random sample of [3,337] articles mentioning the words shortage, scarcity, bottleneck and rationing—the ‘shortage’ words—in conjunction with one or more words indicating economics terms—such as the words ‘economy,’ ‘market,’ or ‘commerce’ (the set  $\mathcal{B}$ ). The ‘shortage’ words above are those more frequently associated with economically-relevant constraints. The inclusion of at

least one economics-related word in the search reduces the likelihood of false positives.<sup>1</sup>

In practice, we construct a list of the 1,000 most popular collocates within five words of the ‘shortage’ words above, and select from this list words that are good candidates for highlighting shortages in particular sectors of the economy or shortages of particular goods. For instance, the most popular words in the list (excluding stopwords) are oil, water, war, time, coal, days, food, cars, people, government, million, labor, state, home, steel, fuel. After removing low-information words such as time or days, as well as other words that convey little content about the topic at hand, such as people or government, we group the remaining terms into four categories that indicate economically-relevant shortages. For ease of exposition, we group the categories into food, industry, labor, and energy, and summarize the resulting search query in Table 1. An article contributes to the shortages index—the set  $\mathcal{S}$ —when two conditions are met: first, any of the shortages words must appear within five words ( $N/5$ ) of one of the topic words; second, the article must contain at least one instance of the economics words. The number of shortage articles is then the sum of the articles in each of the four categories in that period. This definition allows us to weight articles which discuss multiple types of shortages more heavily in the overall index. In the validation section below, we show that requiring the shortage words to appear in close proximity of an actual good of which there is shortage improving the accuracy of the search, and reduces the number of false positives.

The classification into four topics is also supported by the Latent Dirichlet Allocation (LDA) analysis that we perform ex post on a sample of articles satisfying our criterion for inclusion in the index. LDA is a popular unsupervised machine learning technique used for topic modeling in natural language processing, and is designed to discover hidden topics within a collection of documents by analyzing the co-occurrence of words in these documents. There are two inputs to the algorithm. The first is the corpus of documents of text to be analyzed. Here, the corpus is a sample of [13,623] abstracts of newspaper articles mentioning shortages and satisfying our criterion for inclusion in the index.<sup>2</sup> The second input is the number of topics that the LDA analysis should use: we pick four predefined topics, but let the analysis uncover which ones.

The results of the LDA analysis are illustrated and summarized in Figure 1. The most

---

<sup>1</sup> Other potential synonyms, such as “lack” or “paucity” or “insufficiency,” either have a wider range of meanings, or are less likely to be exclusively associated with economic shortages.

<sup>2</sup> As usual, before using the words in the LDA analysis we remove stopwords and numbers. We also stem the words reducing them to a common root.

recurring words for each topic are presented as word clouds in the panels to the left. Topic 1 focuses on energy, topic 2 on water, food and agricultural products, topic 3 on industrial products such as coal, steel, railroads, cars, and topic 4 on jobs. The stacked bar chart to the right visualizes the topic mixtures for each abstract by year (for a total of [13,623] bars, sorted by year). Each article straddles across different topics. Early in the sample, industry-related and, to a lesser extent, food and water shortages dominate the conversation. Energy-related shortages are the most recurring topic throughout most of the 1970s. Labor-related shortages become more prevalent in the post-pandemic years.

## 2.2 Shortages In History

We now present the shortage index, examining spikes and considering the historical context in which they occurred.

Figure 2 plots the shortage index from 1900 through the present. Table 2 lists 30 of the largest spikes in the index throughout history, along with a description of the key events surrounding that episode.<sup>3</sup>

There are several themes behind surges in the shortage index.

- Adverse geopolitical events, especially wars, are associated with shortages. The largest spikes of the index coincide with the end of World Wars I and II.
- Many spikes in the index are associated with labor shortages, especially shortages caused by labor strikes. Shortages caused by strikes are less frequent in the second half of the sample.
- Many post-world war-II jumps in the index are associated with oil and energy shortages, often associated with geopolitical turmoil in the Middle East. The index spikes during the Suez Crisis, during the oil shocks of the 1970s, and around the Iraq invasion of Kuwait in 1990.

---

<sup>3</sup> To calculate spikes, we first calculate the residuals of a regression of the shortage index  $i_t$  against its value two months before and two months after, and look for candidate spikes with the largest values of the residuals. We eliminate a candidate spike if it occurs in a month when the index itself is not a local maximum over a period spanning six months before and six months after (this step ensures that the identified spikes are not just large residuals, but also locally peak values). The observations with the 30 largest values of the residuals are reported in the Table.

- The index spikes again around the global pandemic. The first spike corresponds to shortages of medical equipment and nursing staff at the onset of the Covid-19 pandemic. The second large spike occurs at the beginning of 2022 when global supply bottlenecks are associated with the reopening of many economies following the prolonged mobility restrictions in 2020 and 2021.

### 3 How Accurate is the Shortage Index

We conduct two separate exercises to assess the accuracy of the index. In the first, we verify that our index accurately measures shortages by checking the actual content of the newspaper articles mentioning shortages and verifying that they mention concerns about shortages. In the second, we verify that the index is aligned with alternative measures of shortages for the period in which these alternatives exist and overlap with ours.

#### 3.1 Validation of the Shortage Index

We verify that our index measures accurately shortages by minimizing Type-I and Type-II error. We sample the abstract of 872 articles belonging to the shortage set  $\mathcal{S}$ .<sup>4</sup> By construction, each of these articles contains at least one business-related word as well as one mention of scarcities, shortages or bottlenecks in proximity of a topic word such as energy, food, industry and labor. For each article, we extract the first snippet of text that contains references to shortages. We center the snippet around the shortage word, and set its length to 110 characters, drawing inspiration from Twitter’s original 140-character limit, and striking a balance between brevity and computational and cognitive burden. For instance, examples of recent and old snippets in our sample are respectively “[A]lthough demand remains strong. ... the resulting supply shortage of german manufacturing goods could also...”, from 2021; and “[...] men interested in the industries affected by shortage of steel are anxious to see the strike settled,” from 1901.

We then use the Claude AI assistant ([Anthropic, 2024](#)) to determine whether each snippet mentions current or prospective shortages, rationing, scarcity, or bottlenecks related to goods,

---

<sup>4</sup> Our search query to calculate the index relies on the six U.S. newspapers listed in Section 2. For technical reasons, abstracts are not available for these newspapers since 2015. For the period 2015-2023, we sample the abstract we used for the validation from a larger set of newspapers, including some U.K., Canadian, and Australian ones. A detailed list of the 872 abstracts and sources used for validation is available upon request from the authors. Abstract are short portions of text often containing the initial sentences or the initial paragraph of the entire article.

labor, materials, food, or water. Claude was instructed to return a table of results coding articles either as 1—shortage mentioned—, or zero—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 the coding of each snippet.

Before the classification, we provided Claude with some examples of how we would code the snippets, and made sure that the training sample included false positive, mentioning for instance lack or end of shortages.<sup>5</sup> Use of AI for validation is not foolproof, but we found Claude did as good as 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.<sup>6</sup>

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>7</sup>

Finally, we confirm that restricting the search to include shortage words in proximity of words indicating goods, labor, food or energy improves the accuracy of the search. The share

---

<sup>5</sup> 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 perspective shortages or rationing or scarcity or bottlenecks related to goods, labor, materials, food, water? Just return a table with yes=1, no=0, unsure=99, and a brief explanation. For instance. Article 1 mentions perspective shortages since it mentions that steel shortages will prevail in the near future, so it is a 1. Article 2 says steel shortages caused a plant closure, so it is coded 1. Article 3 says shortage of cars is crimping coal production, so 1. Article 4 mention shortage of cars, so 1. Article 10 mention shortage of workspace, so not really a work shortage, so 0. Article 329 says no shortage of cars has been experienced, so 0.”

<sup>6</sup> 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.” 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.

<sup>7</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.”

of articles mentioning economically-relevant shortages in the set that allows for, but does not require the presence of shortage words is 84.2 percent, corresponding to a Type I error of 15.8 percent, much larger than in our preferred search query. False positives included in this set—that are not captured by our preferred search query—include articles mentioning shortages of political campaign funds, lack of good baseball photos, legislative bottlenecks, and shortage of sunshine.

### 3.2 Comparison with Other Indicators of Supply Constraints

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

In Figure 5, we plot our shortage index, standardized, alongside the New York Fed Global Supply Chain Pressure (GSCPI) Index and the Supplier Delivery Index (SDI). The GSCPI is published by the Federal Reserve Bank of New York and is designed to measure supply chain conditions around the world, using data on both manufacturing and transportation costs. The SDI is published by the Institute of Supply Management (ISM), and captures the result of a monthly survey which asks firms whether they are experiencing longer or shorter wait times compared to the previous month. The value of the SDI is the share of respondents reporting longer wait times plus half the share of respondents reporting no change.

An advantage of our index relative to these two indicators is its availability over a much longer period of time.

Figure 5 shows the GSCPI from its inception in 1998 to the present. Over this period, the correlation between our index and the GSCPI is 0.72. 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.

There is a lower correlation between our index and the SDI. This comparison begins in 1976, when the SDI starts. Our measure has a correlation of 0.37 with the SDI. Both measures spike around the 1979 oil crisis and Covid-19.

Add additional comparison with other indicators?

## 4 Shortages, Economic Activity and Inflation

In this section, we show that increases in shortages are associated with a decline in economic activity and a raise in inflation, the typical effects of supply-side disruptions.



## 4.1 Data and Empirical Model

We estimate vector autoregressions (VAR) models on two dataset. The first model is estimated on annual data from 1900 through 2022. Specifically, We construct world aggregates for real GDP and inflation using data that cover 44 advanced and emerging economies.<sup>8</sup> Data on inflation are from the IMF’s International Financial Statistics and is extended back to 1900 using historical data from [Jordà et al. \(2017\)](#) and, for countries not in their database, [Reinhart and Rogoff \(2009\)](#). To minimize the impact of hyper-inflationary episodes, we winsorize the inflation data at the 1st and 97.5th percentiles. Real GDP per capita data are from [Barro and Ursúa \(2012\)](#), extended through 2022 using the World Bank’s World Development Indicators (WDI) database. Global aggregates are the sum of country variables weighted by real GDP. The use the logarithm of the shortage index and of real GDP, and we use a quadratic trend to detrend log real GDP.

The second model is estimated on U.S. monthly data from January 1971 through December 2023. The dataset includes the shortage index, industrial production, average hourly earnings, and three consumer price indexes (CPI): all items, energy, and food. We transform all series by taking their logarithm and—except for the shortage index—by taking the twelve-month change.

We consider as experiment a one-standard deviation shock to the shortage index. In the annual model, we identify the shock through a Choleski decomposition that orders the shortage index first. Thus, any contemporaneous correlation observed between the shortage index and economic variables reflects the effects of other shocks. At annual frequency, we impose this conservative identification to remove any potential endogeneity of the index to current economic conditions. By contrast, in the monthly model, we isolate the shock through a Choleski decomposition that orders the shortage shock first. This ordering imposes that, on impact, any contemporaneous correlation observed between the shortage index and economic variables reflects the effects of shortage shocks. At monthly frequency, we are less concerned about contemporaneous endogeneity of the shortage index, while we want to explore the potential for immediate effects of shortages. Finally, the annual VAR model is estimated with

---

<sup>8</sup> Advanced economies include Australia, Belgium, Canada, Denmark, Finland, France, Germany, Italy, Japan, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom, and the United States. Emerging market economies include Argentina, Brazil, Chile, China, Colombia, Egypt, Hong Kong, Hungary, India, Indonesia, Israel, Malaysia, Mexico, Peru, Philippines, Poland, Russia, Saudi Arabia, South Africa, South Korea, Taiwan, Thailand, Tunisia, Turkey, Ukraine, Venezuela, and Vietnam.

a constant and two lags, while the monthly VAR includes three lags.

## 4.2 Results

Figure 8 shows the impulse responses to a one-standard-deviation shortage shock over a 10-year horizon estimated for the annual global model. The median response is represented by a solid line, with dark and light shaded areas denoting the 70 and 90 percent confidence intervals, respectively. Following a spike in shortages, we observe a jump in global inflation one year after the shock, followed by a gradual decline over about three years. Shortages induce a decline in the level of global GDP, which peaks at around 0.4 percent before slowly climbing up to pre-shock levels.

Figure 11 shows the impulse responses to a one-standard-deviation shortage shock over a 48-month horizon estimated for the monthly US model. In response to a spike in shortages, there is a sizeable drop in industrial production within two months, followed by a steady return to pre-shock levels that takes about four years. Earnings go up persistently, peaking two years after the shock.

The bottom row of the figure compares the effects of the shock for the three different measures of CPI inflation included in the model. While the contour of the responses is similar, inflation increases, there are also some notable differences. Total CPI inflation peaks at 0.15 percentage points about one year after the shock, and remains above its pre-shock level after four years. The peak in the energy component of CPI is four times as large but less persistent. Finally, the food component of CPI peaks immediately after the shock, well-before the other two measures, returning to its baseline within two years.

All told, the observed negative co-movement between inflation and real activity suggests that shortages induce sizeable and statistically significant adverse supply-side effects. This result applies to the global economy over the past 120 years, and to the US economy over the past 50 years.

## 4.3 Robustness

We consider two robustness exercises.

First, we explore whether the economic effects of shortages vary by category. To this end, we replace the shortage index with the three categorical indexes—focusing on goods, labor,

and commodities—in the baseline annual and monthly models, including them one at the time.

Figure 10 shows the results. The response of global inflation and activity to shocks to the three indexes shares similar profiles. However, the response to labor shortages is less persistent relative to the baseline measure of shortages, especially for economic activity, while the response to goods and commodity shortages is more persistent. For the United States, following labor shortages, IP drops for less than six months, while CPI inflation does not respond. By contrast, goods and commodity shortages exert a larger effect on both IP and CPI relative to the baseline measure.

Second, we confirm that our results are robust to two important modelling choices. The cyan-circled line in Figure 10 shows that results are very similar in a model specification that includes 6 lags instead of 3. The use of fewer lags in the baseline model follows the observation that longer lags are not precisely estimated. The red-squared line shows results for the alternative Cholesky identification that orders the shortage index last in the model, implying that a shortage shock cannot have any contemporaneous effect on the model variables. The responses of inflation to the shock are remarkably similar to the baseline model. The response of IP remains negative but it is more muted relative to the baseline identification, as this model specification attributes the contemporaneous negative correlation between IP and shortages to other shocks in the system.

## 5 Conclusions

Concerns about supply limitations and resource constraints have been a very common theme in press coverage over the past century, whether during wartime, periods of economic upheaval, or as a result of natural disasters or trade disruptions. The specific markets impacted by shortages have evolved with changes in technology and consumption patterns, but the fundamental economic challenges of matching limited supplies with variable demand have remained a persistent source of public anxiety and policy debates. Policymakers and industries have responded to shortages with a variety of coping strategies including rationing essential goods, developing substitutes and alternatives, investing in expanded productive capacity, and promoting conservation. The recurrence of shortage events across decades shows the difficulty of permanently eliminating vulnerabilities in the face of population growth, rising living standards, geopolitical instability and environmental pressures. This paper has presented

a monthly index of shortages going back to 1900. We hope that this analysis can prove useful to researchers.

## 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.
- BARRO, R. J. AND J. F. URSÚA (2012): “Rare Macroeconomic Disasters,” *Annual Review of Economics*, 4, 83–109.
- 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,” .
- 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.
- JORDÀ, O., M. SCHULARICK, AND A. M. TAYLOR (2017): “Macrofinancial History and the New Business Cycle Facts,” *NBER Macroeconomics Annual 2016*, 213–263.
- LAMONT, O. (1997): “Do” shortages” cause inflation?” in *Reducing Inflation: Motivation and Strategy*, University of Chicago Press, 281–306.
- PITSCHNER, S. (2022): “Supply chain disruptions and labor shortages: COVID in perspective,” *Economics Letters*, 221, 110895.

REINHART, C. M. AND K. S. ROGOFF (2009): “This time is different,” in *This Time Is Different*, princeton university press.

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

Table 2: Largest Shortage Surprises

Month	Index	Surprise	Event
Jan-1903	166	118	Nationwide coal shortages
Dec-1906	182	108	Shortage of coal and freight cars in Midwest
Nov-1916	225	106	Nationwide coal shortages
Jan-1918	552	319	Fuel and coal shortages
Dec-1919	286	115	Fuel and coal shortages due to war, strikes
Jul-1920	303	97	Freight car shortage affects coal and steel transportation
Aug-1922	268	135	Coal shortage due to strikes
Aug-1930	89	53	Drought leads to food and water shortages
Jul-1934	128	71	Strike by Teamsters unions in the West Coast
Sep-1939	145	76	Steel shortage due to the beginning of WW2
Aug-1941	493	138	War-related energy, materials and labor shortages
Jan-1943	1036	203	War-related oil, labor and food shortages
Jan-1945	538	174	War-related widespread shortages
Aug-1945	531	144	Labor shortages at the end of war
May-1946	563	180	Strikes by coal workers and fuel shortages
Jan-1948	439	167	Metal, fuel and food shortages
Feb-1950	216	96	Coal shortages amid strikes
Jan-1951	273	102	Labor shortages due to demand from defense industries
Jan-1952	259	120	Nationwide and worldwide shortages
Dec-1956	164	75	Oil shortages due to Suez crisis
Dec-1973	1036	471	Gasoline shortages due to 1973 oil crisis
Jan-1975	255	84	Concerns about gasoline rationing
Feb-1977	403	284	Carter's appeal on energy conservation
Feb-1978	151	66	Concerns about energy shortages
May-1979	553	201	Concerns about energy shortages
Aug-1981	146	49	Gasoline shortages due to 1979 oil crisis
Aug-1990	153	61	Concerns about energy shortages
Apr-2020	284	191	Medical shortages due to COVID-19 pandemic
Jan-2022	529	112	Labor shortages
Oct-2023	205	51	Food shortages in Gaza

Note:

Table 3: Sample Validation

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:



17



Note:

Figure 2: The Shortage Index

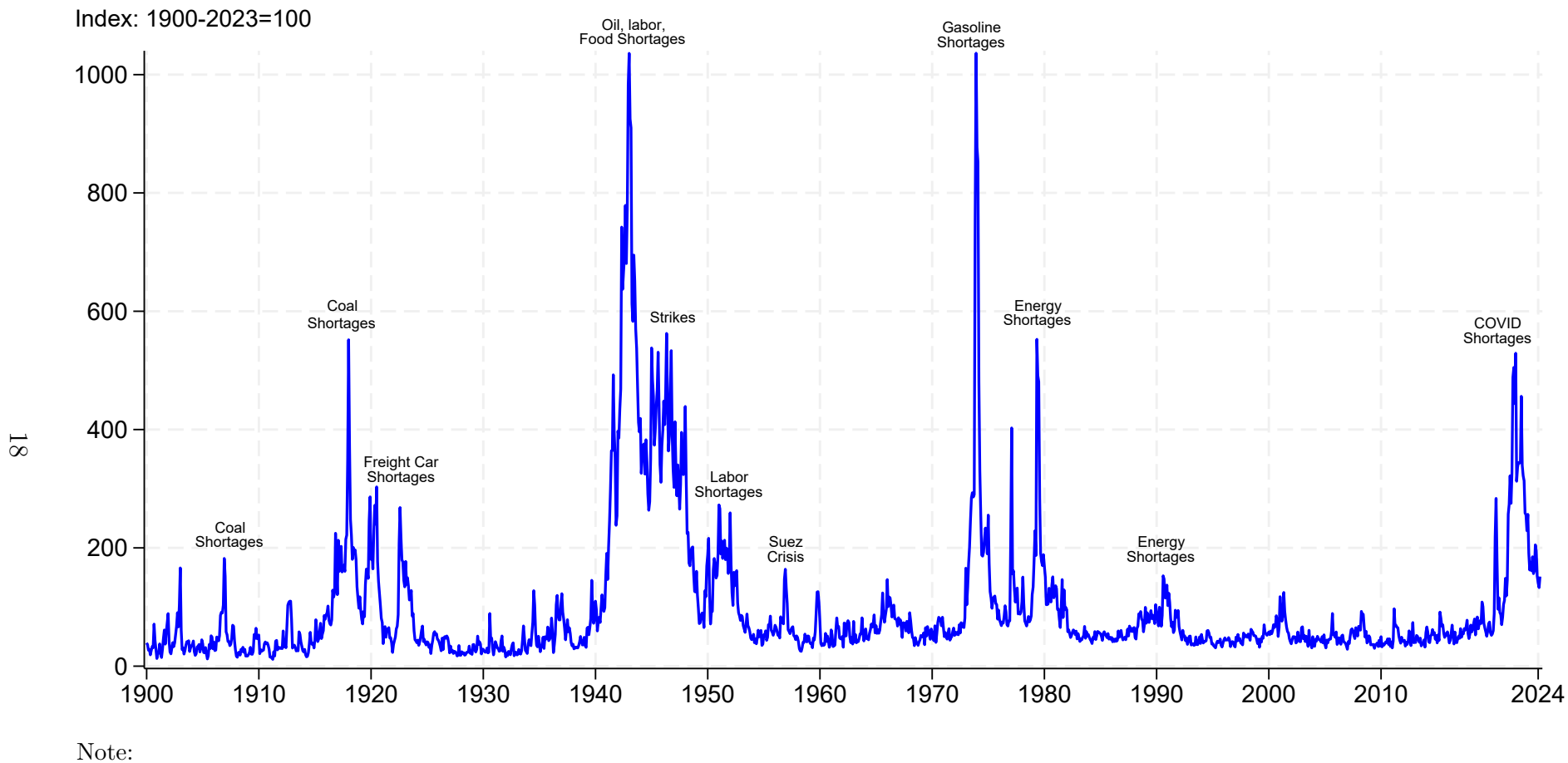
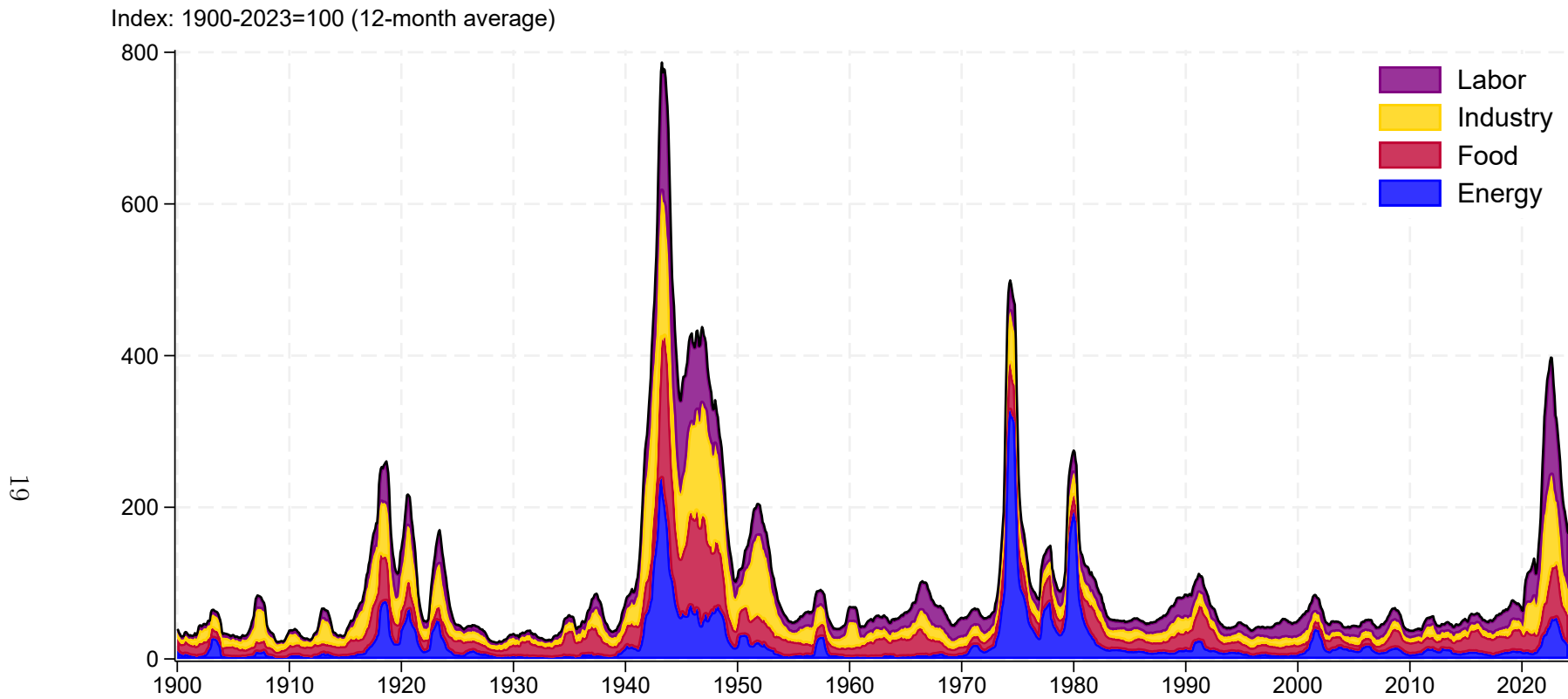
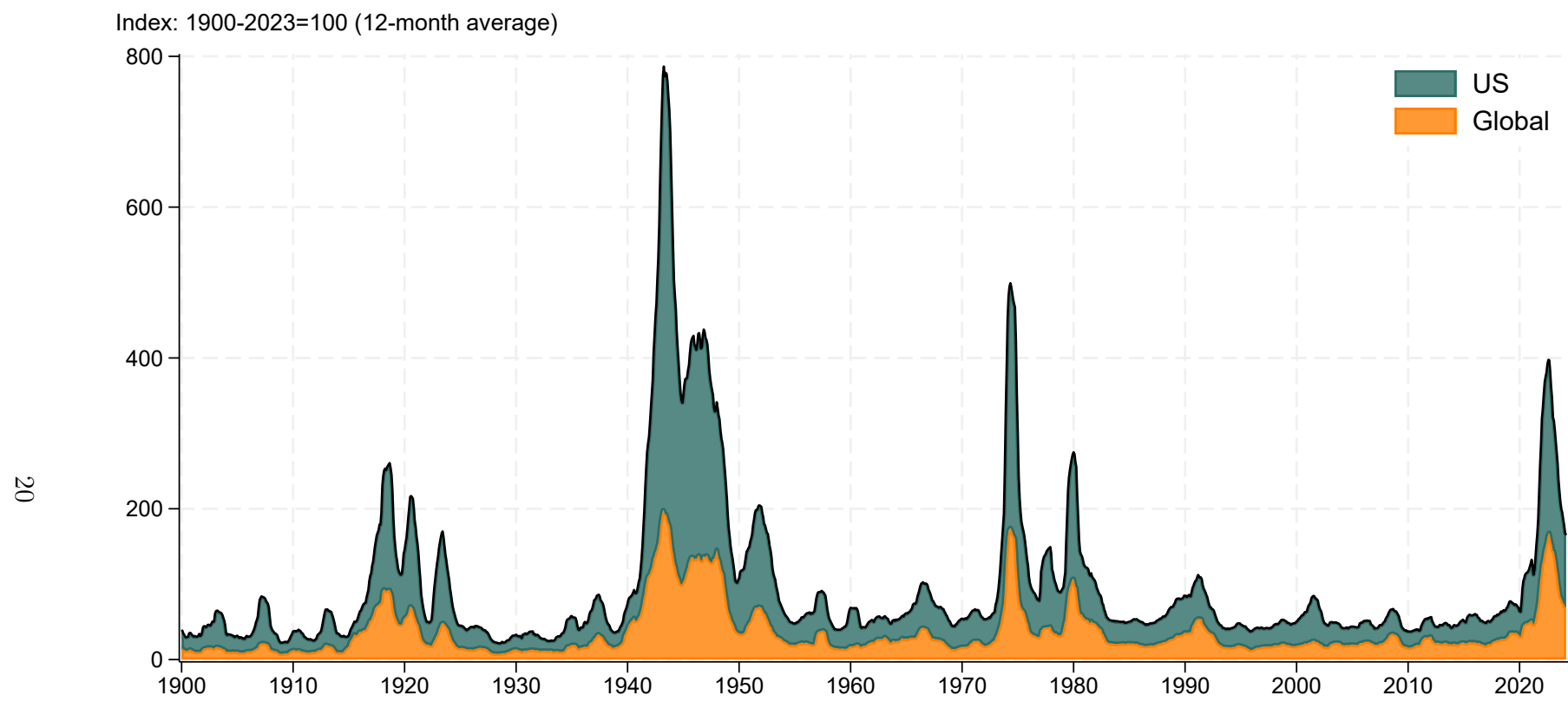


Figure 3: Decomposition by Category



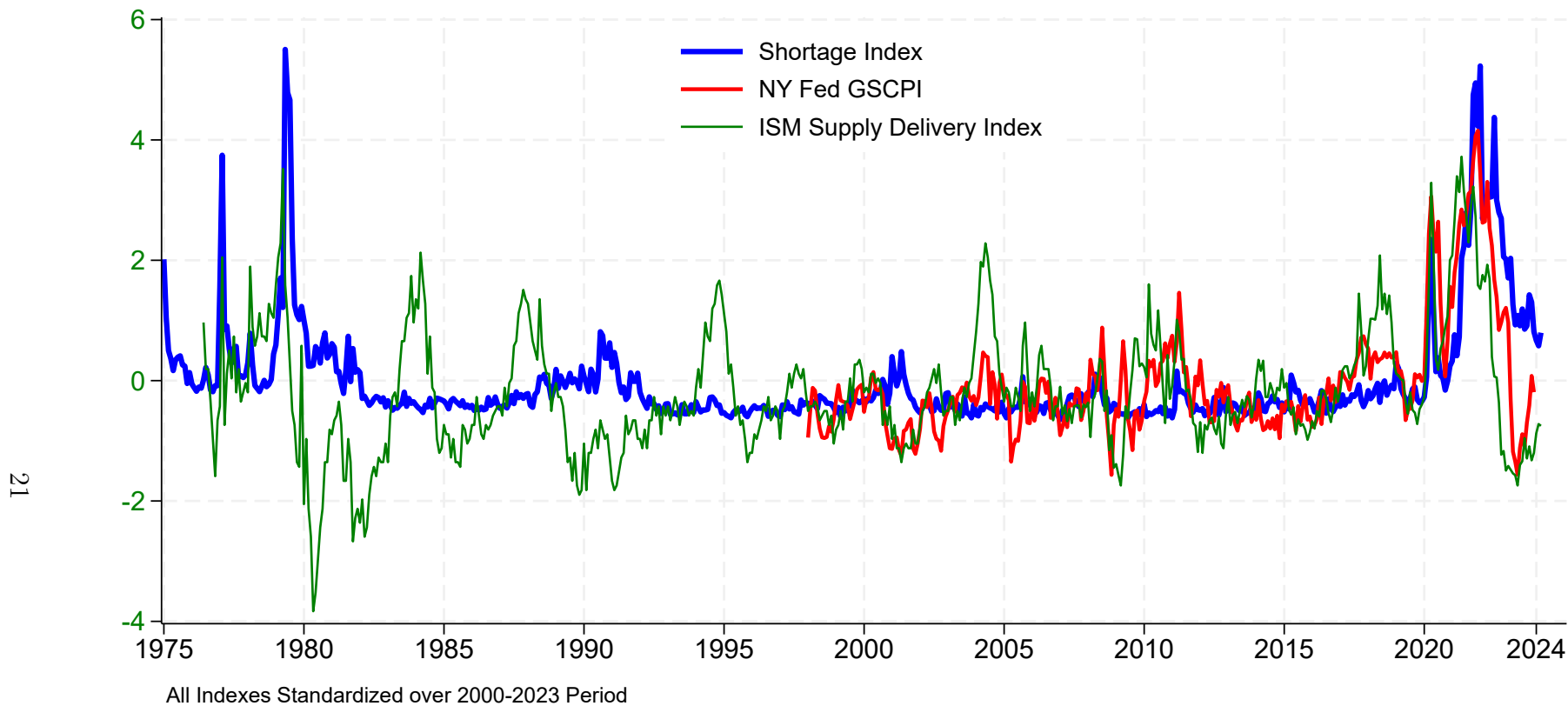
Note:

Figure 4: U.S. vs. Global Shortages



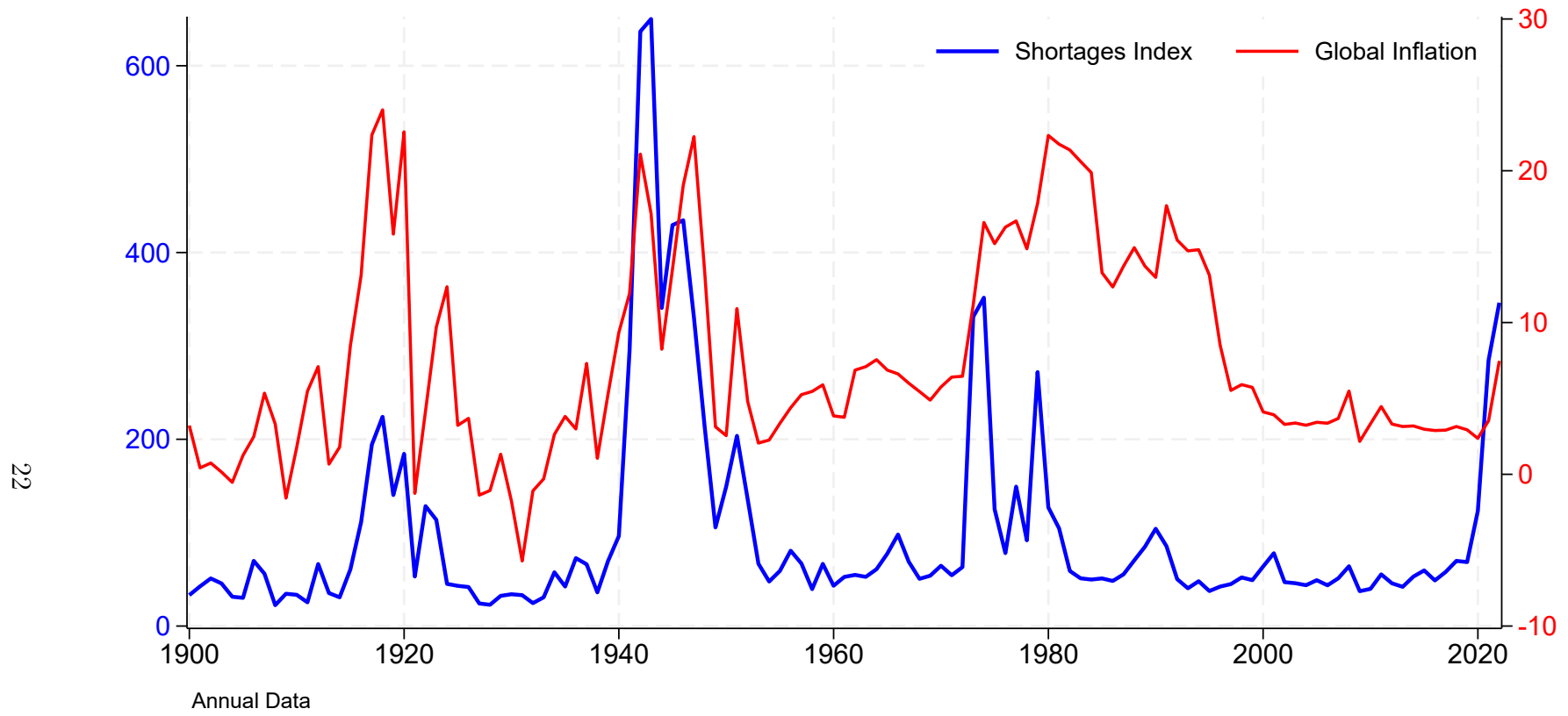
Note:

Figure 5: Comparison to Other Shortage Measures



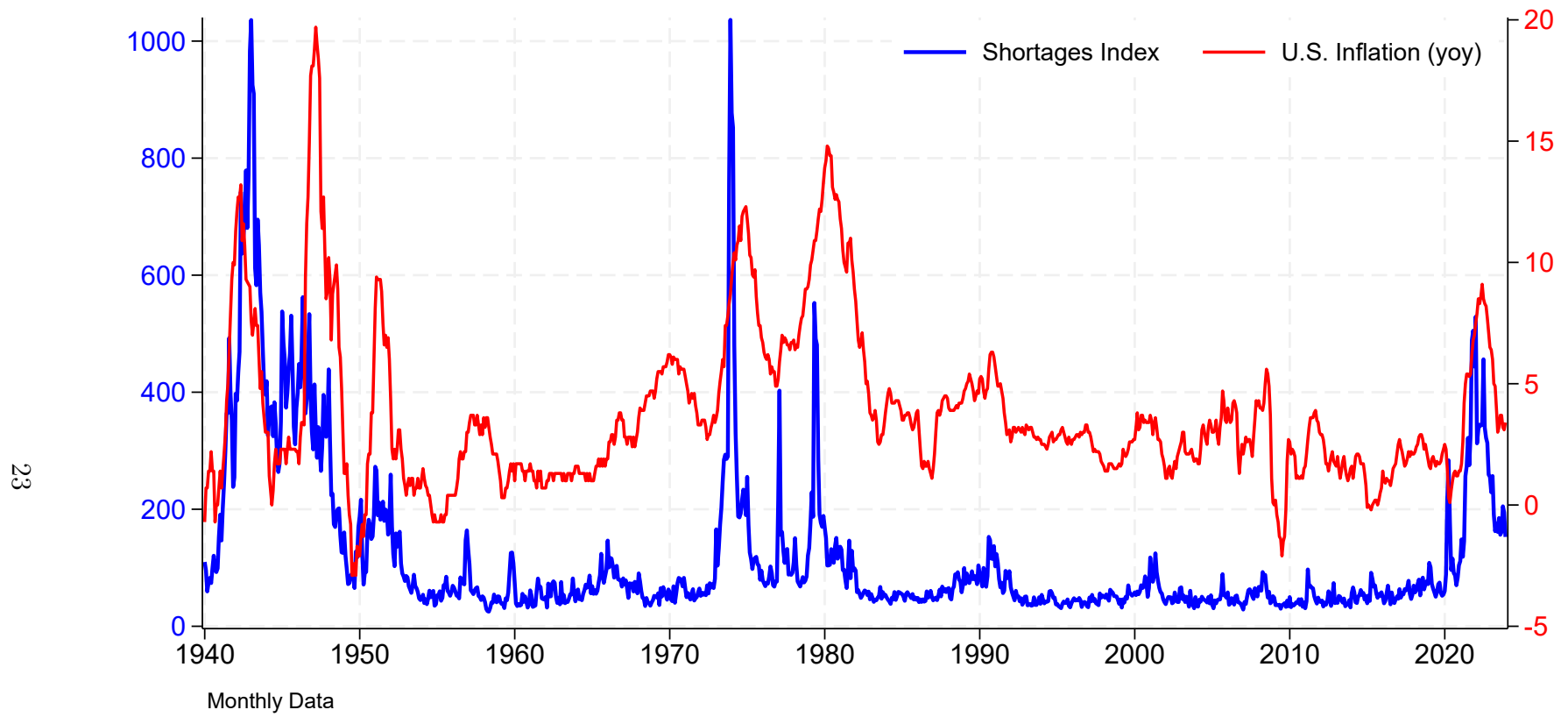
Note: The SDI is computed as the share of respondents reporting longer delivery times plus half the share of respondents reporting no change in delivery times.

Figure 6: Shortages and Global Inflation



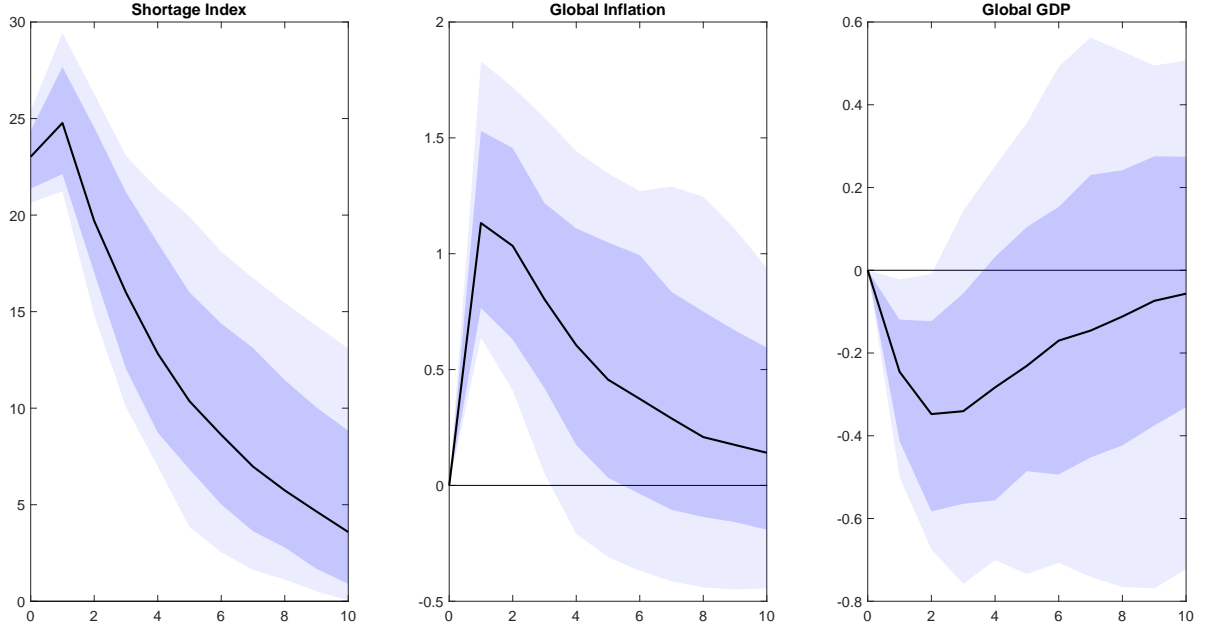
Note:

Figure 7: Shortages and U.S. Inflation



Note:

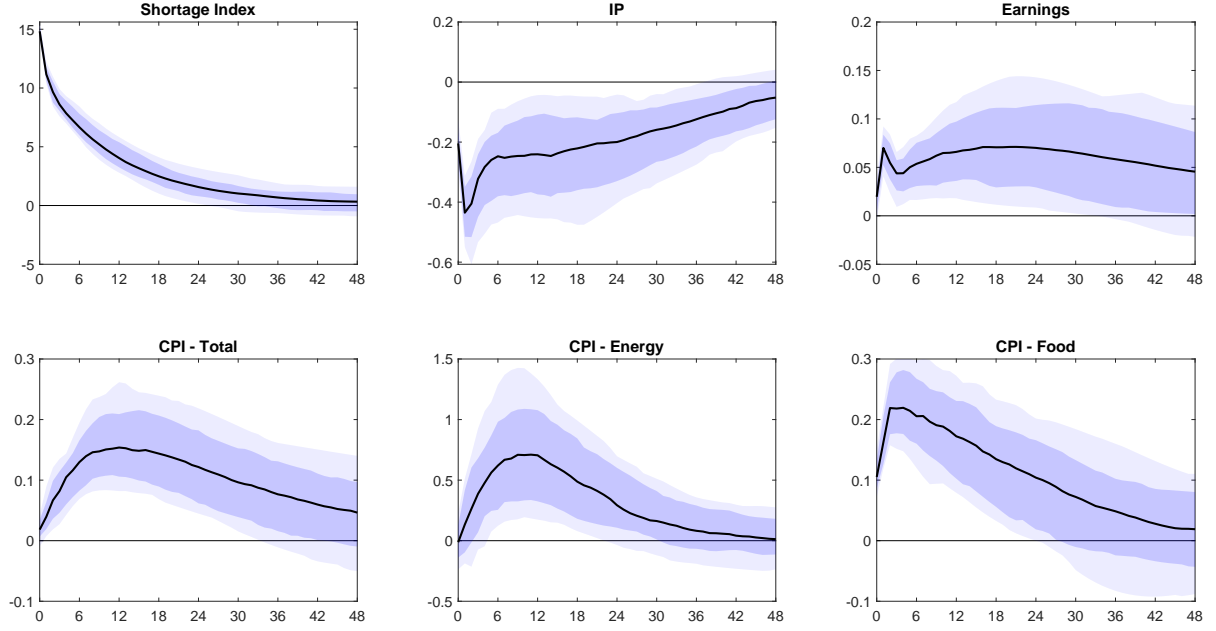
Figure 8: Effects of Shortages on Global Activity and Inflation



Note: The figure plots impulse responses to a one-standard deviation shock to the shortage index. The VAR model is estimated on annual data for the global economy from 1900 through 2021 and includes measures of the shortage index, inflation, GDP, and trade. The solid black lines in the figure plot the central estimates. The dark and light shaded areas denote the 70 and 90 percent confidence intervals, respectively.

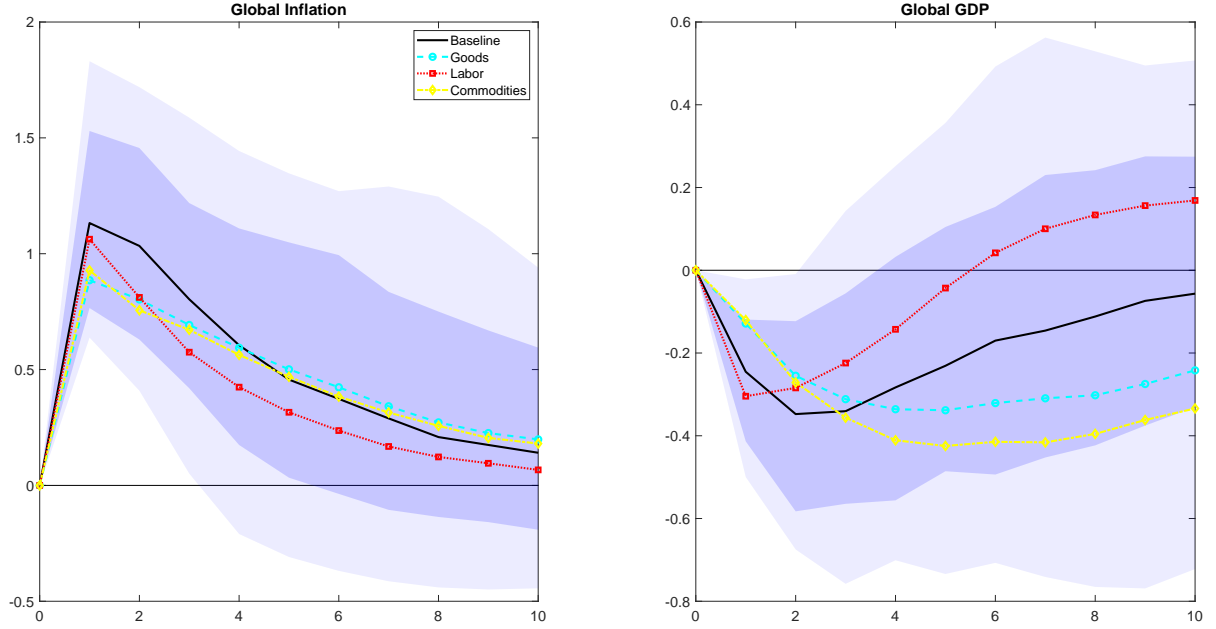


Figure 9: Effects of Shortages on US Activity and Inflation

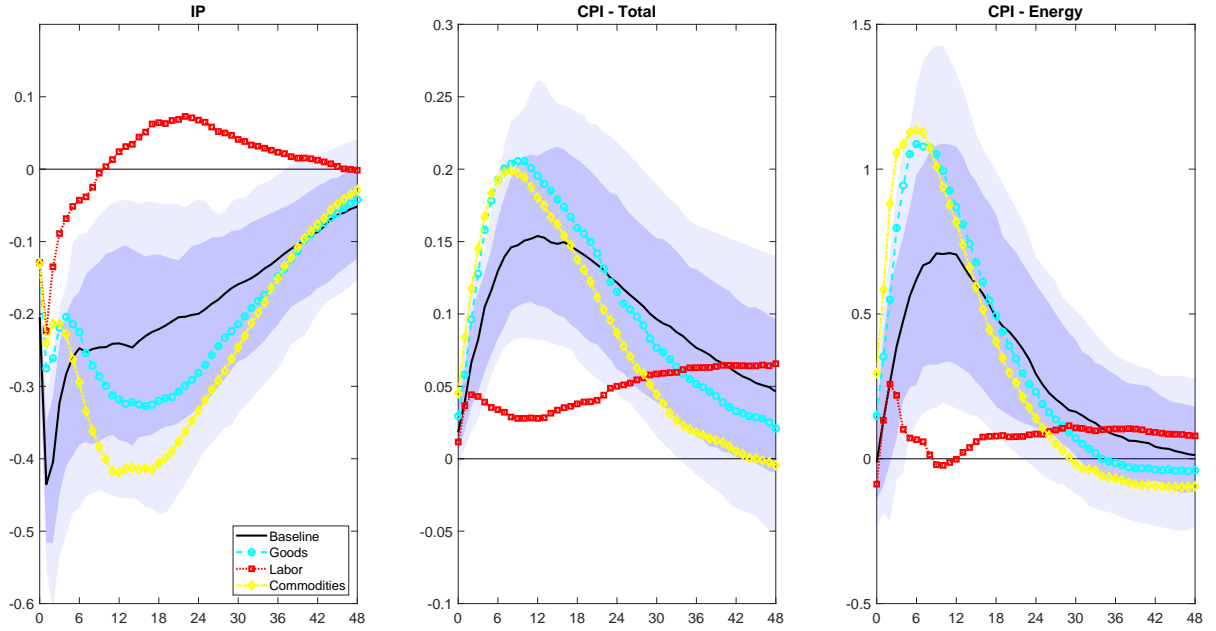


Note: The figure plots impulse responses to a one-standard deviation shock to the shortage index. The VAR model is estimated on monthly data for the United States from 1971 through 2023 and includes measures of the shortage index, industrial production, average hourly earnings, CPI inflation, and the energy and food components of CPI inflation. The solid black lines in the figure plot the central estimates. The dark and light shaded areas denote the 70 and 90 percent confidence intervals, respectively.

Figure 10: Effects of Shortages by Category on Activity and Inflation



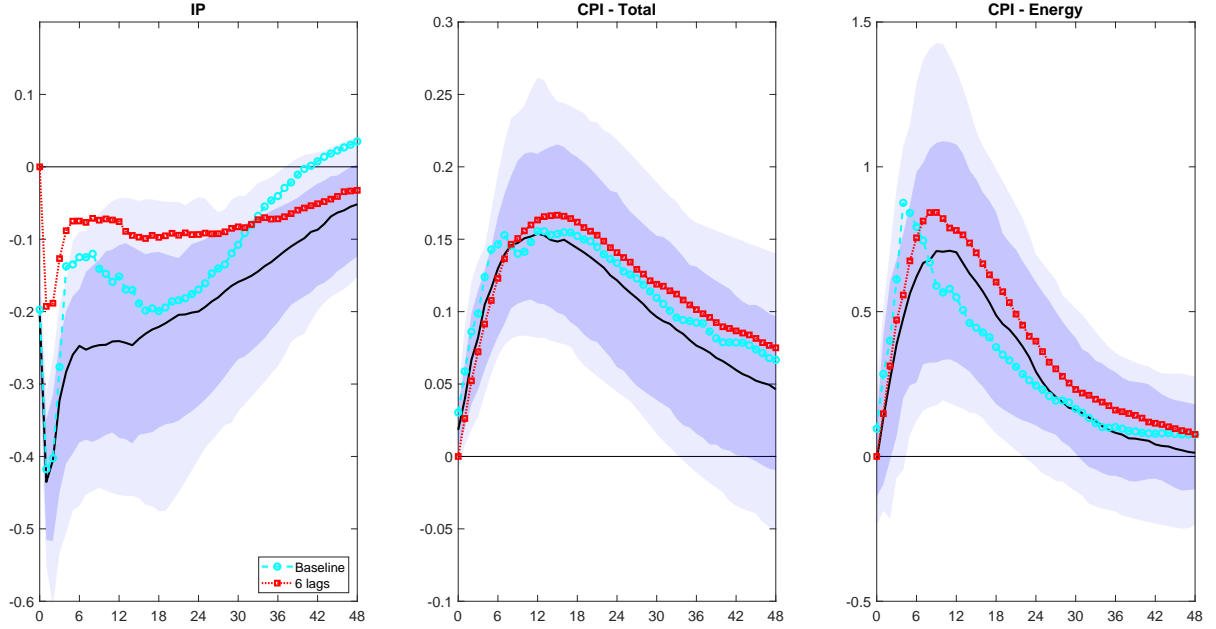
(a) Global Model - Annual Data



(b) U.S. Model - Monthly Data

Note: The figure plots impulse responses to a one-standard deviation shortage shock for selected variables. The top panel depicts results for the VAR model estimated on global annual, while the bottom panel for the VAR model estimated on U.S. monthly data. The solid black lines plot the central estimates, while the dark and light shaded areas denote the 70 and 90 percent confidence intervals, respectively, from the baseline models. The remaining lines plot the central estimates from models where the overall shortage index is replaced by the shortage index for a specific category.

Figure 11: Robustness Checks for the U.S. Monthly VAR



Note: The figure plots impulse responses to a one-standard deviation shock to the shortage index for selected variables. The VAR model is estimated on monthly data for the United States from 1971 through 2023. The solid black lines depict central estimates, while the dark and light shaded areas denote the 70 and 90 percent confidence intervals, respectively, from the baseline models. The remaining lines plot central estimates for alternative specifications of the model.