Team Persistent or Team Transitory? Sectoral Linkage and Inflation Persistence*

Liugang Sheng †
CUHK

Zhentao Shi ‡
CUHK

Steve Pak Yeung Wu §
UC San Diego

March 6, 2024

Preliminary and incomplete

Abstract

This paper develops a high-dimensional vector autoregression framework with unobserved common factors regulated by the Lasso method to explore spillovers across sectors and their implications for aggregate inflation persistence in the US. We demonstrate that accounting for sectoral spillover increases aggregate inflation persistence, particularly during high inflation episodes. A shock to a sector with low persistence in its own inflation could induce persistent aggregate inflation, primarily attributable to cross-sector spillover. Empirically, 13 out of 16 sectors in personal consumption expenditure categories contribute more to aggregate inflation persistence than to their own inflation persistence.

Key words: Inflation, inflation persistence, inflation dynamics, sectoral shocks, sectoral spillover, input-output linkage
JEL code: E3, E5

*Excellent research assistance from Shu Shen is greatly ackowledged.
†Email: lsheng@cuhk.edu.hk
‡The Chinese University of Hong Kong Email: zhentao.shi@cuhk.edu.hk
§UC San Diego. Email: stevepywu@gmail.com

1 Introduction

The rapid surge in inflation in the United States and the global economy post-2020 has had a significant impact on the welfare of consumers. A pivotal factor influencing how policymakers should respond to this inflationary uptick is the persistence of inflation, sparking a series of policy and academic debates. The "Team Persistent" expresses concern that the substantial fiscal stimulus, coupled with accommodative monetary conditions, may lead to elevated and enduring inflation (Blanchard, Summers, among others). In contrast, the "Team Transitory," including Krugman, counter this view by arguing that the inflationary pressure is likely to be temporary or mild.

One of the paramount economic lessons drawn from the COVID-19 pandemic is the crucial role played by sectoral disruptions and spillovers across sectors in the propagation of shocks to the aggregate economy. The significance of sector connections has been underscored by theoretical New Keynesian models presented in Guerrieri et al. (2022), Rubbo (2023), and Afrouzi and Bhattarai (2023). However, a suitable empirical framework has yet to be developed due to the curse of dimensionality and the omission of unobserved common factors, when the standard vector autoregression (VAR) model is applied to the sectoral inflation data. More specifically, the first challenge arises empirically as finer sector categorization may encounter difficulties in estimating the high-dimensional autoregressive coefficient matrix, making statistical inferences challenging. Secondly, omitting the common shocks would lead to biased estimation, pseudo-persistence in sectoral inflation, and spurious spillover between sectoral inflation. For example, a high correlation between sectoral inflation may be caused by a common factor, and a high persistence of sectional inflation may be driven by a persistent common factor.

In this paper, we shed light on the discussion about inflation persistence through the lens of a high-dimensional VAR model augmented with unobserved common factors. Utilizing this framework that incorporates all sectoral inflation data from the United States, we empirically estimate the persistence and spillovers in sectoral inflation, and the implications for aggregate inflation. The approach of estimating interconnectedness empirically allows for more flexible spillover across sectors than imposing the input-output structure, which has been commonly adopted in the literature. The core results of the paper highlight the importance of sectoral spillover for understanding aggregate inflation persistence.

Intuitively, when a variable follows an autoregressive process, the autoregression coefficients possess information about the variable's persistence. Considering a set of sectoral inflations following a vector autoregressive process, the diagonal elements of the autoregression coefficient matrix offer insights into the persistence generated by their own lags.

Importantly, the off-diagonal elements convey information about the empirically estimated spillover from one sector to another. Mathematically, the persistence of aggregate inflation hinges on the largest eigenvalue of the autoregression coefficient matrix. Inter-sector spillover could potentially amplify aggregate inflation persistence.

Empirically estimating the autoregressive coefficient matrix of quarterly PCE sectoral inflation in the United States from 1959-2023 for 16 sectors, we derive four key results. To address the high-dimensional nature of the VAR system, we adopt the Lasso method and estimate it by using the state-of-the-art techniques developed by Miao et al. (2023), which enable the feasible estimation of a high-dimensional VAR model augmented with common factors that permit robust cross-sectional dependence.

Firstly, there are meaningful spillovers across sectors, with many off-diagonal matrix elements estimated to be non-zero. A sector's inflation can induce persistent aggregate inflation through two channels. It may be that the sector itself has high own persistence, as seen in "Housing and Utilities". Alternatively, a sector with relatively low own persistence, such as "Food Services and Accommodations", may exhibit numerous spillovers to other sectors with higher own persistence. It may also be a combination of both channels operating, where a sector has high persistence and spills over into many sectors, as observed in the "Healthcare" sector.

Secondly, aggregate inflation is estimated to be more persistent with spillovers. Without accounting for spillover, it takes about four quarters for the aggregate inflation response to a common factor shock to diminish to one-tenth of the shock size. With spillover considered, this period extends to ten quarters. In terms of the response to sector shocks, spillover-driven inflation persistence contributes more than own inflation persistence to the overall aggregate inflation persistence in 13 out of 16 sectors in PCE categories.

Thirdly, sector spillover gains prominence in high inflation periods. When the off-diagonal elements of the autoregressive matrix are turned off, the predicted inflation value is significantly lower than the realized inflation value during high inflation periods, such as the 1970s to early 1980s and 2021-2023. This indicates that considering sectoral spillover is particularly pertinent in high inflation periods, such as the post-COVID-19 inflation.

The recent surge in inflation in the US has sparked discussions about its causes. Gagliardone and Gertler (2023) attribute the inflation to a combination of an oil shock and easy monetary policy. Benigno and Eggertsson (2023), along with Bernanke and Blanchard (2023), demonstrate that supply-side effects and tightness in the labor market contribute to the recent inflation surge. We contribute to the recent literature that focuses on production networks and incorporate it into considerations of inflation and monetary policy. Recent examples that formulate production network implications for aggregate fluctuations include

Gabaix (2011), Acemoglu et al. (2012), Baqaee and Farhi (2019), and Guerrieri et al. (2022). Rubbo (2023) proposes a monetary policy framework for an economy with an input-output structure and generalizes the New Keynesian framework. Afrouzi and Bhattarai (2023)focus on the theoretical implications of production networks for inflation persistence. Minton and Wheaton (2023)provide empirical support for the dynamic propagation of commodity price shocks through production networks and price stickiness. We complement this literature by providing estimates of an empirical spillover relationship across sectors, which can go beyond the relationships of the input-output network and highlight the spillover implications to inflation persistence.

Layout. We first provide a data description in section 2. Section 3 describes the estimation procedure and the estimated coefficients. Section 4 applies the econometric model to inflation data and provides an analysis to inflation persistence.

2 Data Description

Table 1: Summary Statistics of Sectorial Inflation Rates

Sector	Weight	Mean	Std.	Min	Max
Durable Goods					
Motor vehicles and parts	5.15%	0.56	1.09	-1.55	6.74
Furnishings and durable household equipment	3.44%	0.27	0.87	-1.58	3.83
Recreational goods and vehicles	2.89%	-0.52	1.04	-3.02	2.21
Other durable goods	1.54%	0.42	1.09	-1.72	6.97
Nondurable Goods					
Food and beverages purchased for off-premises consumption	11.30%	0.80	0.99	-1.01	4.64
Clothing and footwear	5.09%	0.27	0.86	-5.91	3.06
Gasoline and other energy goods	3.58%	1.06	6.78	-41.77	19.91
Other nondurable goods	7.86%	0.77	0.71	-0.53	4.71
Services					
Housing and utilities	17.86%	0.94	0.60	-0.06	2.90
Health care	12.01%	1.18	0.78	-0.33	3.50
Transportation services	3.18%	0.92	1.00	-2.58	6.03
Recreation services	3.07%	0.89	0.47	-0.32	2.87
Food services and accommodations	6.45%	1.02	0.64	-0.32	3.81
Financial services and insurance	6.29%	0.99	1.38	-4.09	7.41
Other services	8.15%	0.92	0.60	-0.18	3.25
Final consumption expenditures of nonprofit institutions serving households (NPISHs)	2.15%	0.40	1.51	-4.12	5.26

Notes: All sectors have 258 observations of inflation.

In this empirical application, we apply our econometric methods to the US sectoral inflation. The dataset comprises quarter-on-quarter changes in sixteen sectoral seasonal-adjusted personal consumption expenditure (PCE) price indices. The Bureau of Economic Analysis reports those sixteen disaggregated sectoral price indices (at the lowest level) in the NIPA Table 2.3.4, covering four durable goods sectors, four nondurable goods sectors, and eight service sectors. The sample's raw data consists of quarterly observations spanning from 1959Q1 to 2023Q3. Inflation is measured in log percentage points at a quarterly rate,

$$y_{it} = \inf_{it} = \left(\log \mathrm{PCE}_{it}^{\text{index}} - \log \mathrm{PCE}_{it-1}^{\text{index}}\right) \times 100\%.$$

The sectoral expenditure share weights can be computed using the nominal PCE values in NIPA table 2.3.5. For simplicity, we assume for the empirical analysis that the sector weights are constant across the sample period and equal to the average of the quarterly weights,

$$w_i = \frac{1}{T}\sum_{t=1}^{T} w_{it} \times 100\% = \frac{1}{T}\sum_{t=1}^{T} \frac{\mathrm{PCE}_{it}^{\text{value}}}{\mathrm{PCE}_{t}^{\text{value}}} \times 100\%.$$

The summary statistics of sectoral inflation of 16 sectors and their expenditure share weights are given in Table 1. The three sectors with the largest weight among them are "Housing and Utilities" (17.86%), "Health Care" (12.01%), and "Food and Beverages Purchased for Off-premises Consumption" (11.30%). Over the span of the last sixty years, the average inflation in the services sector has exceeded that of the goods sector. The three sectors, i.e., health care, food services, and financial services have highly advanced average inflation rates — they are all near to or even higher than 1%. The "Health Care" sector, whose weight is the second largest, has the highest average inflation rate of the 16 PCE sectors, reaching 1.18%. Inflation in the durable goods sectors is typically modest, and it has even gone negative for recreational goods and vehicles. The sector of "Gasoline and Other Energy Goods" is unique among nondurable goods since it experiences exceptionally high volatility (6.78%) and average inflation above one percent. Furthermore, when compared to other industries, NPISHs, and financial services have relatively substantial inflation variations.

3 Econometric Models and Estimation

3.1 High-dimensional VAR Model with Common Factors

Let y_{it} be the inflation of the i-th sector at time t, and the N sectors are stacked into $\mathbf{Y}_t = (y_{it})_{i=1}^N$. The VAR is one of the most popular models for multivariate time series. We write a VAR(p)

$$\mathbf{Y}_t = \sum_{k=1}^{p} \mathbf{A}_k \mathbf{Y}_{t-k} + \boldsymbol{\mu} + \mathbf{e}_t, \quad t = 1, 2 \ldots, T \tag{1}$$

where $\boldsymbol{\mu} = (\mu_i)_{i=1}^N$ is the individual specific fixed effects, $\mathbf{A}_k = (a_{ij,k})_{i,j\in[N]}$ is $N \times N$ coefficient matrix, and $\mathbf{e}_t = (e_{it})_{i=1}^N$ is the vector unobservable error term. To capture potential common shocks, we follow Altissimo et al. (2009) and Mayoral (2013) to allow factor structures to enter the error term via

$$\mathbf{e}_t = \boldsymbol{\Lambda}\mathbf{f}_t + \mathbf{u}_t,$$

where $\mathbf{f}_t = (f_{lt})_{l=1}^r$ consists of r latent factors, the $N \times r$ matrix $\boldsymbol{\Lambda}$ stores the factor loadings, and $\mathbf{u}_t = (u_{it})_{i=1}^N$ is the vector of idiosyncratic shocks. To identify the VAR coefficients $\mathbf{A}_k$ for $k \in [p]$, we assume $E[\mathbf{u}_t|\mathbf{Y}_{t-1}, \ldots, \mathbf{Y}_{t-p}, \mathbf{f}_t, \boldsymbol{\mu}] = 0$. Without loss of generality, we normalize $\boldsymbol{\Lambda}'\boldsymbol{\Lambda}/N = \mathbf{I}_r$ to identify both $\boldsymbol{\Lambda}$ and $\mathbf{f}_t$.

Compared to textbook VAR models, our model includes unobserved common factors and allows the factor loading to vary across sectors. Common factors are used to capture unobserved monetary policy, fiscal policy, inflation expectations, TFP shocks, news shocks, etc. Factor loading reflects the contemporary sensitivity of sectoral inflation in response to common factors. Previous studies adopting the VAR framework to model the dynamics of sectoral inflation usually did not include common factors in the model and thus suffered criticism that omitting the common shocks would lead to biased estimation, pseudo-persistence in sectoral inflation, and spurious correlation between sectoral inflation. For example, the high correlation between sectoral inflation may be caused by a common factor and high persistence in sectional inflation may be driven by a persistent common factor. Moreover, including common factors allows us to separate aggregate and sectoral shocks, and thus we can explore the impulse responses of sectoral and aggregate inflation to both shocks.

Previous studies on the nexus between sectoral and aggregate inflation also included common factors. Altissimo et al. (2009) and Mayoral (2013) focused on the heterogeneity in the diagonal coefficients a_{ii}, but ignored intersectoral spillover by assuming the off-diagonal coefficients $a_{ij} = 0$ for $j \neq i$. Our model captures the heterogeneity in sectoral persistence and the spillover between sectors through the VAR specification. The diagonal coefficients

a_{ii} in $\mathbf{A}_1$ capture the autoregressiveness of each sector i, and the off-diagonal coefficients $a_{ij}, j \neq i$ capture the impact of the lagged inflation in sector j on the inflation of sector i in period t. This is an important generalization of our model from the literature. Moreover, Altissimo et al. (2009) and Mayoral (2013) also assume the factor loadings for the common factors are the same for all sectors, which are less likely to hold empirically. Our model relaxes this assumption.

3.2 Estimation Procedure

Collect all observable regressors as a Np-dimensional vector $\mathbf{X}_t = \left(\mathbf{Y}'_{t-1}, \mathbf{Y}'_{t-2}, \ldots, \mathbf{Y}'_{t-p}\right)'$. Stack the $N \times T$ matrix $\mathbf{Y} = (\mathbf{Y}_1, \ldots, \mathbf{Y}_T)$ and the $Np \times T$ matrix $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_T)$. The unknown parameters in this model to be estimated is the $r \times T$ matrix $\mathbf{F} = (\mathbf{f}_1, \ldots, \mathbf{f}_T)$, and the $N \times r$ factor loading matrix $\mathbf{\Lambda}$, and the stacked slope coefficients (an $N \times Np$ matrix) $\mathbf{B} = (\mathbf{A}_1, \ldots, \mathbf{A}_p)$, The sum of squared residual is

$$\mathrm{SSR}\left(\mathbf{\Theta}\right) = \left\| \mathbf{Y} - \mathbf{B}\mathbf{X} - \mathbf{1}_N \boldsymbol{\mu}' - \mathbf{\Lambda}\mathbf{F}' \right\|_{\mathcal{F}}^2 ,$$

where $\left\|\cdot\right\|_{\mathcal{F}}$ is the matrix Frobenius norm, and the parameters is $\mathbf{\Theta} = \{\mathbf{B}, \boldsymbol{\mu}, \mathbf{\Lambda}, \mathbf{F}\}$. This is the criterion function that the OLS estimator minimizes.

The data we employed is a panel data containing 257 quarterly time series observations of 16 sectors. As a result, each $\mathbf{A}_k$ has 256 ($= 16 \times 16$) unknown coefficients to be estimated by data. We will need regularization to control the quality of estimation.

The way of regularization counts on the economic context. While the diagonal elements of $\mathbf{A}_k$ characterize the auto-dynamic structure of a time series, the off-diagonal embody the dynamic interactions with other variables. Altissimo et al. (2009) and Mayoral (2013)'s specification is equivalent to a very sparse model by setting $p = 1$ and all $a_{ij,k} = 0$ to shut off such interaction, where "sparsity" here refers to the fact that most elements in $\mathbf{A}_k$ are zeros and only a few (the diagonal elements) are to be determined by the data. It is natural for us to go one step further by keeping a sparse configuration, but allow the data to determine which off-diagonal elements are zero.

Lasso (Tibshirani, 1996) is the most well-known technique to produce a data-driven sparse estimator. Lasso and its variants has also been intensively studied in econometrics in recent years (Kock and Callot, 2015; Su et al., 2016; Lee et al., 2022). Our paper use Lasso to estimate unknown coefficients.

Denote $\mathbf{A}_k^{\circ} = \mathbf{A}_k - \mathrm{diag}\left(\mathbf{A}_k\right) = \left(a_{ij,k}\mathbb{I}\left\{i \neq j\right\}\right)_{i,j\in[N]}$ as a copy $\mathbf{A}_k$ with the diagonal elements suppressed at 0. Similarly, denote $\mathbf{B}^{\circ} = \left(\mathbf{A}_1^{\circ}, \ldots, \mathbf{A}_p^{\circ}\right)$. A Lasso estimator can minimize

$$\min_{\mathbf{\Theta}} \left\{ \frac{1}{NT} \mathrm{SSR}\left(\mathbf{\Theta}\right) + \gamma \left\| \mathbf{B}^{\circ} \right\|_1 \right\}$$

where $\|A\|_1 = \sum_{i,j} |a_{ij}|$ is the L_1-norm of a generic matrix A which spares the diagonal elements from the penalty. However, it is well-known that the Lasso penalty introduces non-trivial bias toward 0, and therefore Miao et al. (2023) propose multi-step estimation procedure based on Lasso to refine it.

3.3 Estimated Autoregressive Coefficient Matrix

In practical estimation, the number of lags p and the number of factors r must be determined; Miao et al. (2023) provide a systematic procedure for these important hyperparameters. This is the method we use for estimation, which we elaborate on in the Appendix. The information criteria select one common factor ($\widehat{r} = 1$) and one vector autoregressive lag as well. Figure 1 presents the estimated values of transition matrix $\widehat{\mathbf{A}}_1$ with a heat map, where the ith row and jth column is the estimated influence of the lagged inflation in sector j on the sector i's current inflation, i.e., $\widehat{a}_{ij}$. Positive and negative effects are represented by the red and blue cells and white cells for no impact.

There are a few noteworthy discoveries. First, 75 of the 256 total elements are non-zero, and the majority of those non-zero estimates are positive. It implies that the Lasso method has reduced many parameters for estimation and thus makes the model more parsimonious. The LASSO shrinkage offers a more parsimonious model and also a conservative estimate of the spillover. Despite the conservative estimate of the non-zero elements, Approximately thirty percent of the elements in $\widehat{\mathbf{A}}_1$ are not zero, indicating that many important impacts across sectors.

Secondly, the diagonal elements capture the impact of each sector's inflation in the previous period on its own inflation in the current period. Our estimation suggests that all of the diagonal elements in $\widehat{\mathbf{A}}_1$ are positive, but the values vary across sectors. It goes without saying that a high level of the diagonal elements will result in persistent sectoral inflation; however, it should be noted that the diagonal elements do not fully capture sectoral inflation persistence due to the spillover effects across sectors. To ease the comparison, we rank the sectors by the estimates of $\widehat{a}_{ii}$ from the largest to the smallest in Table 2. It shows that the sectors with high values of $\widehat{a}_{ii}$ are "Housing and Utilities" (0.88), "Recreational Goods and Vehicles" (0.76), and "Health Care" (0.68), which tend to have high inflation persistence. As a comparison, "Financial Services and Insurance", "Clothing and Footwear", and "Transportation Services" have low values of $\widehat{a}_{ii}$, which tend to have low inflation persistence. Later we will discuss sectoral inflation persistence as a response to common shocks and sectoral

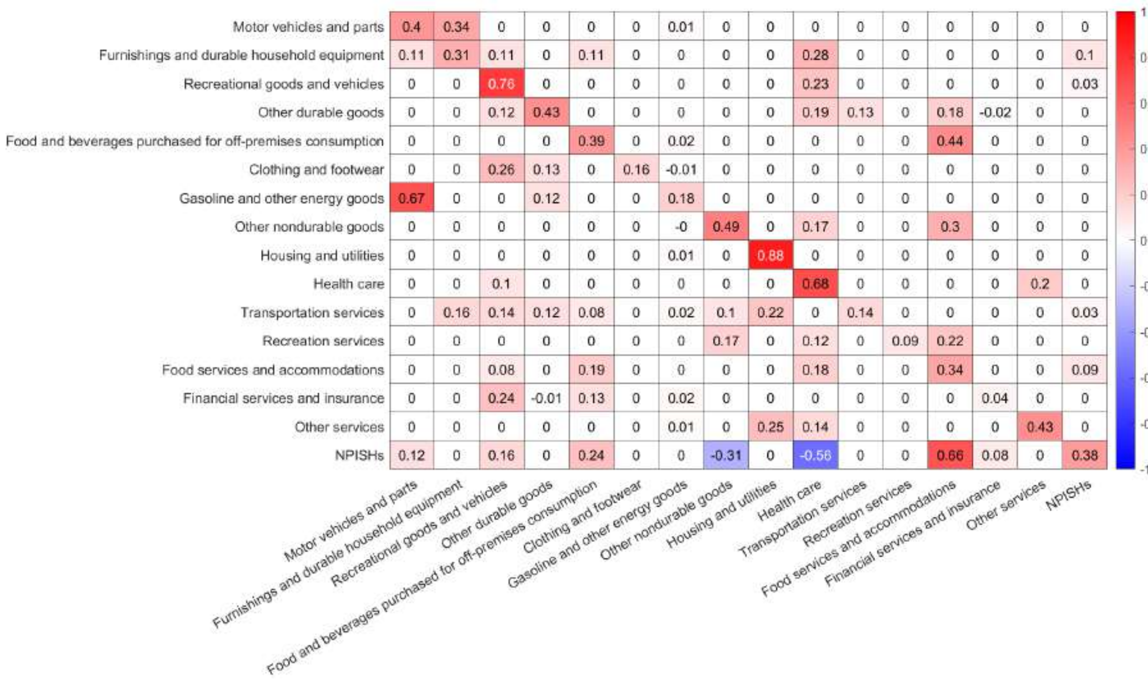

Figure 1: Transition Matrix $\widehat{\mathbf{A}}_1$

Notes: The values of the elements of transition matrix $\widehat{\mathbf{A}}_1$, i.e., $\widehat{a}_{ij}$, are represented by the different color cells. Positive and negative effects are represented by the red and blue cells and white cells for no impact. The value of $\widehat{a}_{ij}$ is shown by the color intensity of every cell.

shocks more formally by using impulse responses.

Third, most of the non-zero off-diagonal elements are positive, indicating a positive inter-sectoral spillover in sectoral inflation. For example, a positive price shock of "Motor Vehicles and Parts" (perhaps due to demand shock) tends to increase the price of gasoline and other energy goods. A positive price shock in "Food Services and Accommodations" also leads to a hike in the price of food and beverages purchased for off-premises in the next period. To better illustrate the sectoral spillover, we also present a directional network graph of sectoral inflation in Figure 2. The value of $\widehat{a}_{ij}$ is shown by the thickness of the connection. Positive and negative effects are represented by the red and blue curves, and sectoral weights are represented by the size of the circle. The arrow indicates the direction of cross-sector spillover. It shows that several sectors significantly influence the inflation in other sectors, such as "Health Care", "Recreational Goods and Vehicles", "Food and Beverages Purchased for Off-premises Consumption", and "Food Services and Accommodations". This result suggests that ignoring the sectoral spillover in inflation may omit an important mechanism for the high persistence of aggregate inflation.

In general, a sector's inflation can radiate to aggregate inflation via two mechanisms:

Table 2: Ranking of Sectoral Inflation Persistence

Sector	$\widehat{a}_{ii}$ ($i=$ Sector)	Sector	$\widehat{a}_{ii}$ ($i=$ Sector)
Housing and utilities	0.88	Final consumption expenditures of non-profit institutions serving households (NPISHs)	0.38
Recreational goods and vehicles	0.76	Food services and accommodations	0.34
Health care	0.68	Furnishings and durable household equipment	0.31
Other nondurable goods	0.49	Gasoline and other energy goods	0.18
Other services	0.43	Recreation services	0.16
Other durable goods	0.43	Transportation services	0.14
Motor vehicles and parts	0.40	Clothing and footwear	0.09
Food and beverages purchased for off-premises consumption	0.39	Financial services and insurance	0.04

the sector's own persistence and sectoral spillover. As shown in the "Housing and Utilities" sector, its high persistence can pass through to aggregate inflation on its own. Certain sectors, like "Food Services and Accommodations", have comparatively low persistence, yet they can have a significant impact on other sectors. Moreover, a sector may concurrently transmit persistence to the aggregate inflation via both pathways. While being highly persistent, the two sectors "Health Care" and "Recreational Goods and Vehicles" also spill over persistence to a multitude of other sectors.

4 Empirical Analysis

4.1 Theoretical Implications

As the empirical strategy has determined $p = 1$ and $r = 1$, we use the VAR(1) model with a single factor

$$\mathbf{Y}_t = \mathbf{A}_1 \mathbf{Y}_{t-1} + \mathbf{e}_t \tag{2}$$
$$\mathbf{e}_t = \mathbf{\Lambda} f_t + \mathbf{u}_t. \tag{3}$$

for the following discussion. For simplicity, we coerce $\mathbf{Y}_t$ in the above expression has zero-mean. To avoid confusing with the individual specific vector $\boldsymbol{\lambda}_i$, we keep using the the upper case $\mathbf{\Lambda}$ here as a $N \times 1$ matrix of factor loadings for the sole common factor.

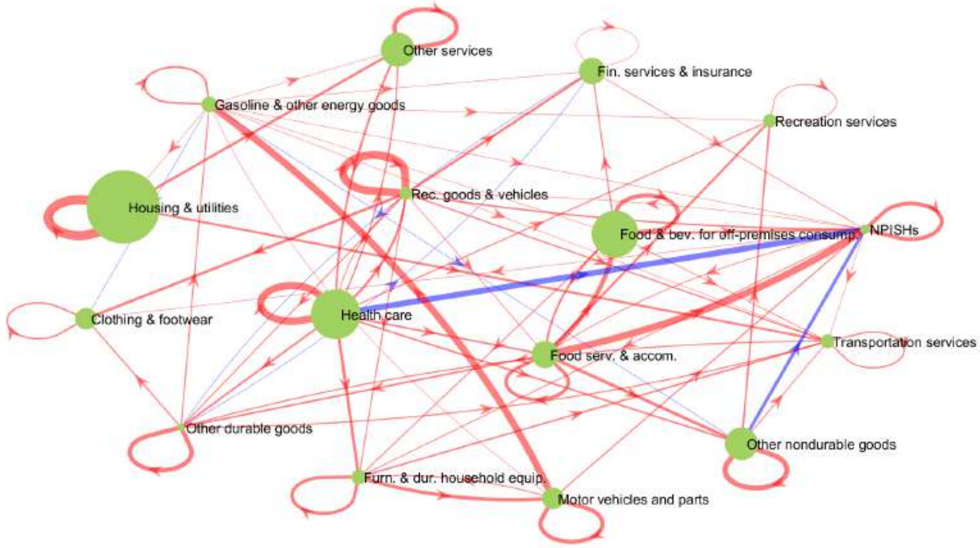

Figure 2: Network Graph of the Transition Matrix $\widehat{\mathbf{A}}_1$

Notes: The value of $\widehat{a}_{ij}$ is shown by the thickness of the connection. Positive and negative effects are represented by the red and blue curves with the arrow. The weights of 16 sectors are represented by the size of the circle. The arrow indicates the direction of cross-sector spillover.

4.1.1 Impulse Responses

When all eigenvalues of $\mathbf{A}_1$ lie in $(-1, 1)$, we can write the VAR(1) as to the moving average representation:

$$\mathbf{Y}_t = (\mathbf{I} - \mathbf{A}_1(\mathbb{L}))^{-1} (\boldsymbol{\Lambda} f_t + \mathbf{u}_t) = \sum_{s=0}^{\infty} \mathbf{A}_1^s (\boldsymbol{\Lambda} f_{t-s} + \mathbf{u}_{t-s}).$$

Define the information set $\mathcal{I}_t = ((\mathbf{u}_s, f_s)_{s=-\infty}^t)$. The conditional expectation is

$$\begin{aligned} \mathbb{E}_t [\mathbf{Y}_{t+h}] &= \mathbb{E} \left[\sum_{s=0}^{\infty} \mathbf{A}_1^s (\boldsymbol{\Lambda} f_{t+h-s} + \mathbf{u}_{t+h-s}) | \mathcal{I}_t \right] \\ &= \sum_{s=0}^{h-1} \mathbf{A}_1^s \mathbb{E} [(\boldsymbol{\Lambda} f_{t+h-s} + \mathbf{u}_{t+h-s}) | \mathcal{I}_t] + \sum_{s=0}^{\infty} \mathbf{A}_1^{h+s} (\boldsymbol{\Lambda} f_{t-s} + \mathbf{u}_{t-s}). \end{aligned} \tag{4}$$

Assume u_t is a martingale difference sequence (m.d.s.) with respect to so that $\mathbb{E} [\mathbf{u}_{t+s} | \mathcal{I}_t] = 0$ for all $s \geq 1$. Regarding f_t, Mayoral (2013) models it as a static factor whereas Altissimo et al. (2009) model it as a dynamic factor.

In our empirical context, we observe that f_t has very low persistence with AR(1) coefficient 0.025. Therefore, we assume that $(\mathbf{u}_t', f_t)$ is m.d.s. to spare the modeling of the dynamics in the factor as well. Under this assumption, the sectoral impulse responses (IRs), defined as the responses of $\mathbf{Y}_t$ to unit changes in $\mathbf{u}_t$ and $\mathbf{f}_t$, have been considered. In view of (4), the h-period ahead IR of sectoral inflation to idiosyncratic sectoral shock $\mathbf{u}_t$ is

$$\mathbf{IR}_{\mathbf{Y}}^{\mathbf{u}}(h) = \partial \mathbb{E}_t\left[\mathbf{Y}_{t+h}\right]/\partial \mathbf{u}_t' = \mathbf{A}_1^h$$

and the h-period ahead IR of sectoral inflation to common factor f_t is

$$\mathbf{IR}_{\mathbf{Y}}^{f}(h) = \partial \mathbb{E}_t\left[\mathbf{Y}_{t+h}\right]/\partial f_t = \mathbf{A}_1^h \boldsymbol{\Lambda}$$

The aggregate inflation is the weighted average of sectoral inflation, i.e., $\pi_t = \mathbf{w}'\mathbf{Y}_t$ is a convex combination, where $\mathbf{w} = (w_i)_{i=1}^N$ are the weights of CPI of N sectors. Thus, the h-period ahead IR of aggregate inflation is also the convex combination of the corresponding sectoral IRF:

$$\mathbf{IR}_{\pi}^{\mathbf{u}}(h) = \partial \mathbb{E}_t\left[\mathbf{w}'\mathbf{Y}_{t+h}\right]/\partial \mathbf{u}_t' = \mathbf{w}'\mathbf{IR}_{\mathbf{Y}}^{\mathbf{u}}(h) = \mathbf{w}'\mathbf{A}_1^h \tag{5}$$

$$\mathbf{IR}_{\pi}^{f}(h) = \partial \mathbb{E}_t\left[\mathbf{w}'\mathbf{Y}_{t+h}\right]/\partial f_t = \mathbf{w}'\mathbf{IR}_{\mathbf{Y}}^{f}(h) = \mathbf{w}'\mathbf{A}_1^h \boldsymbol{\Lambda}. \tag{6}$$

We can further define the percentage contribution of sector i to the aggregate IRF to common shocks at horizon h as follows:

$$\mathbf{PC}_i(h) = \frac{w_i \mathbf{IR}_{\mathbf{Y}_i}^{f}(h)}{\mathbf{IR}_{\pi}^{f}(h)} = \frac{w_i \left(\mathbf{A}_1^h\right)_{i\bullet} \boldsymbol{\Lambda}}{\mathbf{w}'\mathbf{A}_1^h \boldsymbol{\Lambda}} \tag{7}$$

4.1.2 Aggregation

The dynamics of the aggregate inflation $\pi_t = \mathbf{w}'\mathbf{Y}_t$ is AR(1) when $\mathbf{A}_1$ is proportional to an identity matrix in the form diagonal matrix with homogeneous elements a_1 so that

$$\mathbf{Y}_t = a_1 \mathbf{Y}_{t-1} + \mathbf{e}_t.$$

This is the classical panel data model with homogeneous coefficients, and the heterogeneity is embodied by the individual-specific intercept only. For the aggregate inflation, this model implies the simply relationship $\pi_t = a_1 \pi_{t-1} + e_t^{(w)}$, where

$$e_t^w = \mathbf{w}'\mathbf{e}_t = \mathbf{w}'\left(\boldsymbol{\Lambda} f_t + \mathbf{u}_t\right)$$

is orthogonal to π_{t-1}.

The above model is restrictive. For example, if $\mathbf{A}_1$ is a diagonal matrix with heterogeneous elements,

$$\pi_t = \mathbf{w}'\mathbf{A}_1\mathbf{Y}_{t-1} + e_t^{(w)}. \tag{8}$$

It is clear that on the right-hand side the regressors the N-dimensional vector $\mathbf{Y}_{t-1}$ cannot be written as some linear function of the scalar π_{t-1}. Notice that though $\mathbf{A}_1$ only the diagonal elements are zeros, in general the $1 \times N$ vector of $\mathbf{w}'\mathbf{A}_1$ is not a sparse matrix. It is not recommend to use Lasso to directly estimate the N-unknown slope coefficients in (8) that are associated with $\mathbf{Y}_{t-1}$ if N is non-trivial relative to T.

4.1.3 Diagonalizable Transition Matrix

More general, if $\mathbf{A}_1$ is diagonalizable, then it can be decomposed as $\mathbf{A}_1 = \mathbf{V}\mathbf{D}\mathbf{V}^{-1}$, where $\mathbf{D} = (d_j)_{j=1}^N$ is a diagonal matrix of (possibly complex) eigenvalues (order from high to low in terms of each one's modulus), and the $\mathbf{V}$ matrix consists of N columns of eigenvectors. The aggregate inflation is

$$\pi_t = \mathbf{w}'\mathbf{V}\mathbf{D}\mathbf{V}^{-1}\mathbf{Y}_{t-1} + e_t^{(w)}.$$

Regarding the IRF, in this case,

$$\mathbf{IR}_\pi^f(h) = \mathbf{w}'\mathbf{V}\mathbf{D}^h\mathbf{V}^{-1}\mathbf{\Lambda} = \widetilde{\mathbf{w}}'\mathbf{D}^h\widetilde{\mathbf{\Lambda}} = \sum_{j=1}^N \widetilde{w}_j\widetilde{\lambda}_j d_j^h \tag{9}$$

where $\widetilde{\mathbf{w}} = (\widetilde{w}_j)_{j=1}^N = \mathbf{V}'\mathbf{w}$ is the rotated (by $\mathbf{V}$) weight vector, and $\widetilde{\mathbf{\Lambda}} = (\widetilde{\lambda}_j)_{j=1}^N = \mathbf{V}^{-1}\mathbf{\Lambda}$ is the rotated factor loading vector. Notice that $\widetilde{w}_j = \mathbf{V}'_{\bullet j}\mathbf{w}$ is the inner project of the jth eigenvector and $\mathbf{w}$, and similar $\widetilde{\lambda}_j = \mathbf{V}'_{\bullet j}\mathbf{\Lambda}$ is the inner product of the jth eigenvector and $\mathbf{\Lambda}$. Although d_j can be a complex number, any complex eigenvalue always is accompanied by its conjugate and therefore $\mathbf{IR}_\pi^f(h)$ is remains a real number.

In our empirical comparison we normalize $\mathbf{IR}_\pi^f(h=0) = 1$, which is equivalent to

$$\mathbf{IR}_\pi^{f,\text{normalized}}(h) = \frac{\mathbf{IR}_\pi^f(h)}{\mathbf{IR}_\pi^f(0)} = \frac{\mathbf{IR}_\pi^f(h)}{\mathbf{w}'\mathbf{A}_1^0\mathbf{\Lambda}} = \frac{\mathbf{IR}_\pi^f(h)}{\mathbf{w}'\mathbf{\Lambda}} = \frac{\mathbf{IR}_\pi^f(h)}{\widetilde{\mathbf{w}}'\widetilde{\mathbf{\Lambda}}}$$

for $\mathbf{w}'\mathbf{\Lambda} \neq 0$, which holds in general. Given the decomposition (9), we can also write it as

$$\mathbf{IR}_\pi^{f,\text{normalized}}(h) = \frac{\sum_{j=1}^N \widetilde{w}_j\widetilde{\lambda}_j d_j^h}{\widetilde{\mathbf{w}}'\widetilde{\mathbf{\Lambda}}} = \sum_{j=1}^N \delta_j d_j^h$$

where $\delta_j = \frac{\widetilde{w}_j \widetilde{\lambda}_j}{\widetilde{\mathbf{w}}' \widetilde{\mathbf{\Lambda}}}$ is the weight of $\widetilde{w}_j \widetilde{\lambda}_j$ relative to the inner product $\widetilde{\mathbf{w}}' \widetilde{\mathbf{\Lambda}}$, and it is clear that $\sum_{j=1}^{N} \delta_j = 1$. Therefore, $\mathbf{IR}_{\pi}^{f,\text{normalized}}(h)$ can be interpreted as a linear combination of $(d_j^h)_{j=1}^N$, where the weight δ_j is determined by $\mathbf{w}$, $\mathbf{\Lambda}$ and $\mathbf{V}$. For the weights δ_j that are complex, they appear in conjugate pairs and therefore $\mathbf{IR}_{\pi}^{f,\text{normalized}}(h)$ is guaranteed to be a real number.

Once the weight vector $(\delta_j)_{j=1}^N$ is provided, over the horizon h the IRF is determined by the eigenvalues d_j for $j \in [N]$. When h grows the smaller eigenvalues shrinks to 0 at a much faster speed then the larger ones, and therefore the persistence of the IRF is mainly determined by the large eigenvalues if the associated weights are non-trivial. (In our empirical application we computed the weights for each section. All are bounded away from 0.)

This theoretical insight is important as it shows that the heterogeneous sectoral persistence (diagonal entries in $\mathbf{A}_1$) is insufficient to understand the aggregate inflation persistence. The previous literature on the linkage between sectoral and aggregate inflation usually focuses on the heterogeneity in sectoral persistence by assuming a diagonal matrix of $\mathbf{A}_1$, which ignores the sectoral spillover and thus may lead to biased estimation in both the sectoral and aggregate inflation persistence.

Remark 1. When $\mathbf{A}_1$ is diagonalizable, we can rotate the raw vector $\mathbf{Y}_t$ into $\mathbf{V}^{-1}\mathbf{Y}_t$ by premultiply $\mathbf{V}^{-1}$ on both sides of (2) to obtain

$$\begin{aligned} \mathbf{V}^{-1}\mathbf{Y}_t &= \mathbf{D}\mathbf{V}^{-1}\mathbf{Y}_{t-1} + \mathbf{V}^{-1}\mathbf{e}_t \\ \mathbf{V}^{-1}\mathbf{e}_t &= \mathbf{V}^{-1}\boldsymbol{\mu} + \mathbf{V}^{-1}\mathbf{\Lambda} f_t + \mathbf{V}^{-1}\mathbf{u}_t \end{aligned}$$

This system is very similar to Mayoral (2013), where the transition matrix is a diagonal matrix $\mathbf{D}$.

4.2 Intersectoral Spillover and Inflation Shocks

To highlight the importance of spillovers and provide a sense of which sector tops the spillover effect, the h-period ahead prediction of aggregate inflation π_t can be calculated as the sector inflation weighted by the PCE category weight plus the future common factor shocks and sectoral shocks:

$$\mathbb{E}_t\left[\pi_{t+h}\right] = \mathbb{E}_t\left[\boldsymbol{w}^\top \mathbf{Y}_{t+h}\right] = \mathbb{E}_t\left[\boldsymbol{w}^\top \left(\mathbf{A}_1\right)^h \mathbf{Y}_t + \sum_{s=0}^{h-1} \left(\mathbf{A}_1\right)^s \mathbf{\Lambda}\mathbf{f}_{t+h-s} + \mathbf{u}_{t+h-s}\right] = \boldsymbol{w}^\top \left(\mathbf{A}_1\right)^h \mathbf{Y}_t.$$

Given a 1% sector inflation shock to each sector i, the contribution of the shock to the aggregate inflation h-period ahead via its own persistence is given by

$$\text{Own Persistence}_i\,(h) = w_i \left((\mathbf{A}_1)^h \right)_{i,i} \tag{10}$$

and the contribution of the shock to the aggregate inflation h-period ahead via its spillover to other sectors (spillover-driven persistence) can also be obtained by summing the persistence of other sectors that sector i spillover to

$$\text{Spillover-driven Persistence}_i\,(h) = \sum_{j=1, j\neq i}^{N} w_j \left((\mathbf{A}_1)^h \right)_{i,j}. \tag{11}$$

Figure 3 plots the relative contribution to aggregate inflation of the own sector inflation persistence and spillover-driven persistence (sum up to 100%) 1-period later and the infinite sum of all the future periods. Following the impact of sector inflation in six of the sixteen sectors, aggregate inflation in the 1-period prediction surpassed its own driving power driven by spillover effects, the average contribution of which is 68%. Cross-sector spillovers have an even more significant long-term influence. The spillover-driven persistence exceeds 50% in most sector inflation shocks (12 sectors). The same idea is seen in Figure 4, in which we plot the level of inflation contributed by the two channels rather than the share. One period after the 1% shock, the responses of inflation to housing, health care, and food services shocks are significant. The aggregate inflation rises by more than 12 basis points. Notably, Housing and Healthcare have the greatest cumulative inflation effects of 3.5% and 2.3% in the long run and it is largely because of the spillover-driven inflation.

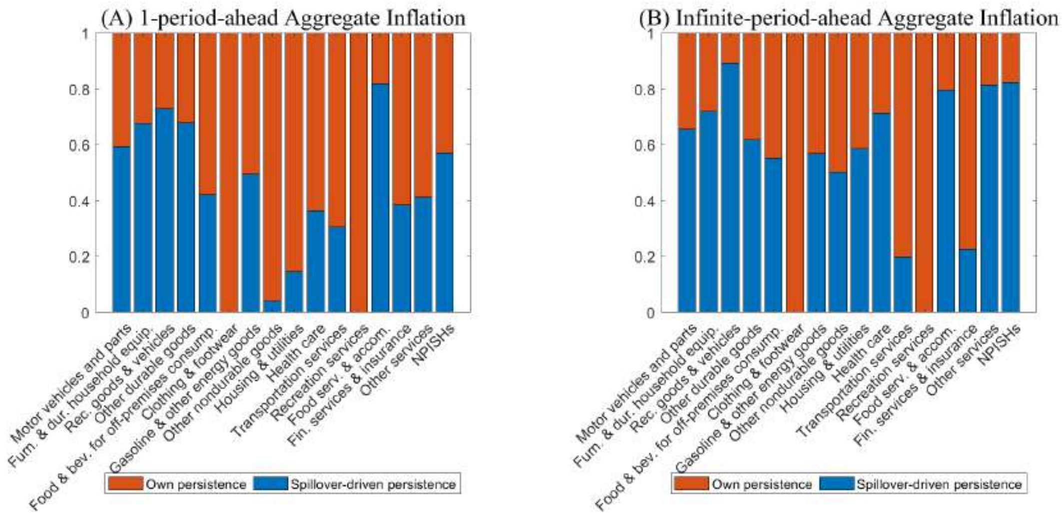

Figure 3: Shares of Own and Spillover-driven Persistence Contributing to 1-period-ahead Aggregate Inflation and the Sum of Infinite-period-ahead Aggregate Inflation (Given 1% Sector Inflation Shock)

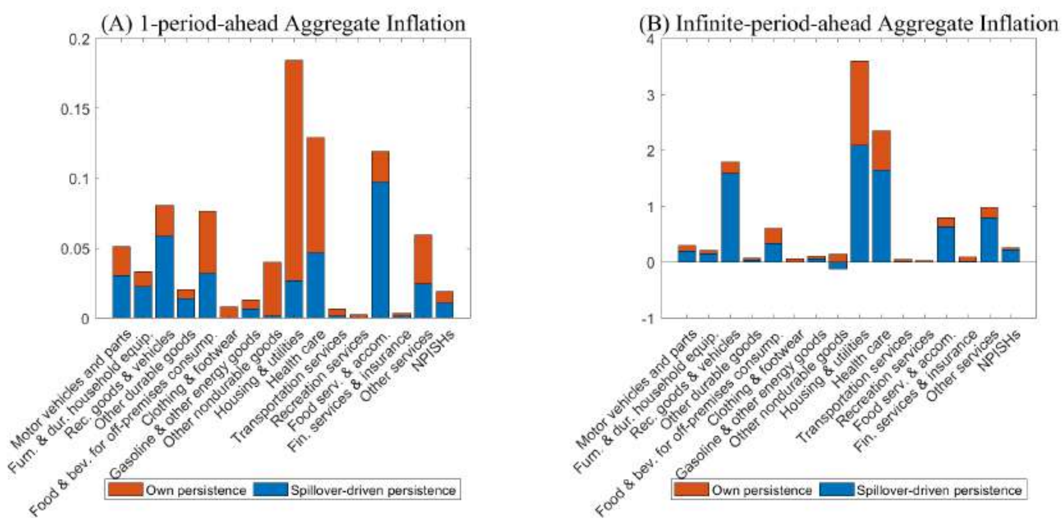

Figure 4: Levels of Own and Spillover-driven Persistence Contributing to 1-period-ahead Aggregate Inflation and the Sum of Infinite-period-ahead Aggregate Inflation (Given 1% Sector Inflation Shock)

4.3 Common Factors

4.3.1 Aggregate Impulse Responses to Common Factors

To highlight the importance of sectoral heterogeneity in persistence and intersectoral spillover, we also estimate two restricted models and compare them with our general model:

(a) Heterogeneous AR coefficients without spillover: $\mathbf{A}_1$ is a diagonal matrix.

(b) Homogeneous AR coefficients without spillover: $\mathbf{A}_1 = a_1$ is a scalar.

The number of common factors is set to one for estimating the two restricted models, in line with the estimates of our general model. Furthermore, the calculation of the impulse response is according to (6) and is consistent for the three models. After estimating the three models, Figure 5 displays the impulsive responses of aggregate inflation to common factors.

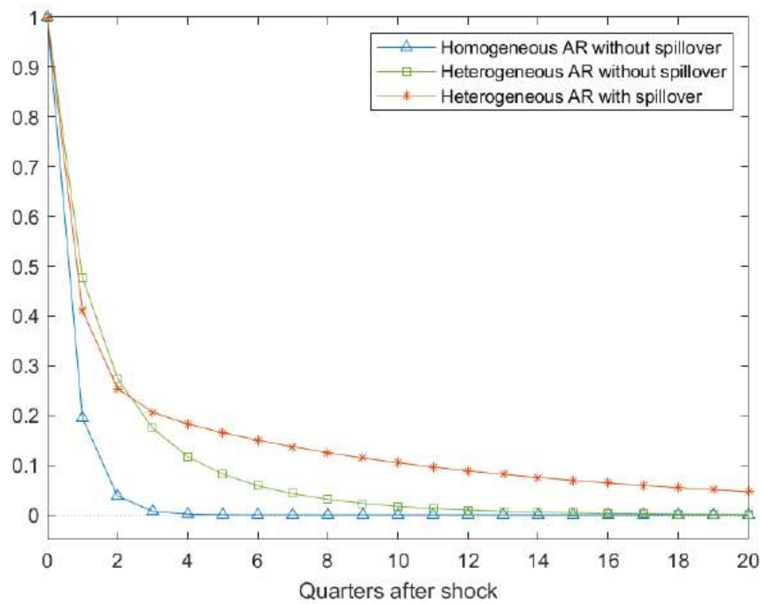

Figure 5: Aggregate Impulse Responses to Common Factors

Notes: Lines marked with upward-pointing triangles, squares, and asterisks represent the aggregate IRs to common factors from the model of homogeneous AR coefficients without spillover, the model of heterogeneous AR coefficients without spillover, and our general model, respectively.

As can be seen from Figure 5, the general model that takes intersectoral spillover into account in this paper can better capture the high persistence of US inflation than the two restricted models. There will have been a 41 basis point increase in aggregate inflation one quarter following a 1% common shock. Remarkably, despite relatively larger reductions in the year after the shock, IRs progressively stabilize and the declines in each quarter thereafter decrease quarter by quarter, which are all below 0.02%. Longer-term effects of common factors are still evident, as seen by the fact that five years after the common shock, the impulse response of aggregate inflation is still 5 basis points.

The performance of these two restricted models is unsatisfactory. The homogenous AR coefficients model without spillover has impulse responses of aggregate inflation to common factors that nearly vanish a year after it occurs with IRs quickly declining to 0 because the model ignores intersectoral spillover and sectoral heterogeneity. In a model that solely accounts for sectoral heterogeneity and ignores intersectoral spillover, the IR of aggregate inflation reduces to one-tenth the size of the shock in about four quarters after facing the common factor shock, which is reached in the tenth quarter when the spillover effects are taken into account. The reduction in IRs increases drastically after that and is insignificant after three years, less than 0.01%.

4.3.2 High and Low Inflation Period

By examining the entire time series of aggregate inflation, we can see that the data variations before and after 1990 differ significantly. The United States' Great Inflation of the 1970s was greatly contributed by the oil crisis and the US inflation rate has been rather steady since then. We estimate the sub-sample below to investigate the model's sensitivity to inflation data. Following structural breakpoint testing, 1990Q4 was identified as the important threshold, and the full sample is split into two subsamples: the high inflation period (1959Q2-1990Q4) and the low inflation period (1991Q1-2020Q4). It should be mentioned that the effects of COVID-19 will cause another bout of severe inflation after 2020, so this period will be momentarily disregarded here.

Figure 6 displays the outcomes of the impulsive responses to common factors in the high and low inflation period. The persistence of the impact of aggregate inflation on the common shock in the United States mainly comes from the Great Inflation in the 1970s. Prior to 1990, the United States inflation rate sustains a rise of 26 basis points after a year following a unit common shock, more than the full sample result, before gradually declining. While the inflation in the low-inflation period has little persistence, there is a brief and notable increase in aggregate inflation one quarter after the common shock, with no difference from zero in the influence of shock thereafter. Moreover, this figure shows that our general model is more sensitive to capturing the persistence of inflation than models that exclude inter-sector spillover effects. The near consistency of the results of the heterogeneous AR coefficients without spillover model throughout the two periods of markedly differing inflation rates points to the significance of inter-sector spillovers.

4.3.3 Sectoral Contribution to Aggregate Inflation Persistence

The quantification of each sector's contribution to aggregate persistence helps to separate the significant roles that different sectors play in inflation persistence. (7) can be used to evaluate the percentage contribution of the sector i to the impulse responses of aggregate inflation to the common factor at horizon h. Figure 7 shows the results of the percentage contribution of each sector to the aggregate inflation persistence for $h = 1, ..., 20$.

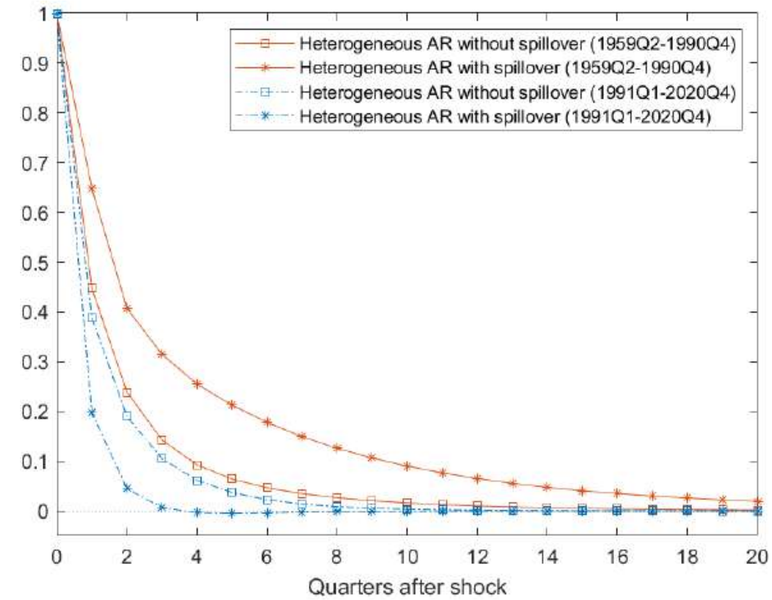

Figure 6: Aggregate Impulse Responses to Common Factors in the High and Low Inflation Period

Notes: Lines marked with squares and asterisks represent the aggregate IRs to common factors from the model of heterogeneous AR coefficients without spillover and our general model, respectively. Solid and dash-dotted lines represent the aggregate IRs to common factors in the high and low inflation periods, respectively.

The distribution of sectoral contributions is related to the length of time since the shock. The sector of "Gasoline and Other Energy Goods" contributes the most to the aggregate IR in the first 3 months after the common shock, peaking at 77%, but a sharp fall afterward. The key driver for the persistence of aggregate inflation is the "Housing and Utilities" sector, the contribution of which stays steady at about 19%, between 6 months and 3 years after the common shock, ranking first. Three years later, the sector of "Health Care" defeats "Housing and Utilities" and becomes the number one contributing sector. Its contribution rate increases instead of decreasing, reaching 18% in the fifth year. Unlike the housing sector, which depends mostly on its own strong persistence, the health sector also contributes to the aggregate inflation through another major channel, which is sector spillover. As the third most persistent sector, "Health Care" will only contribute monotonically decreasingly to aggregate inflation if sectoral spillover effects are ignored.

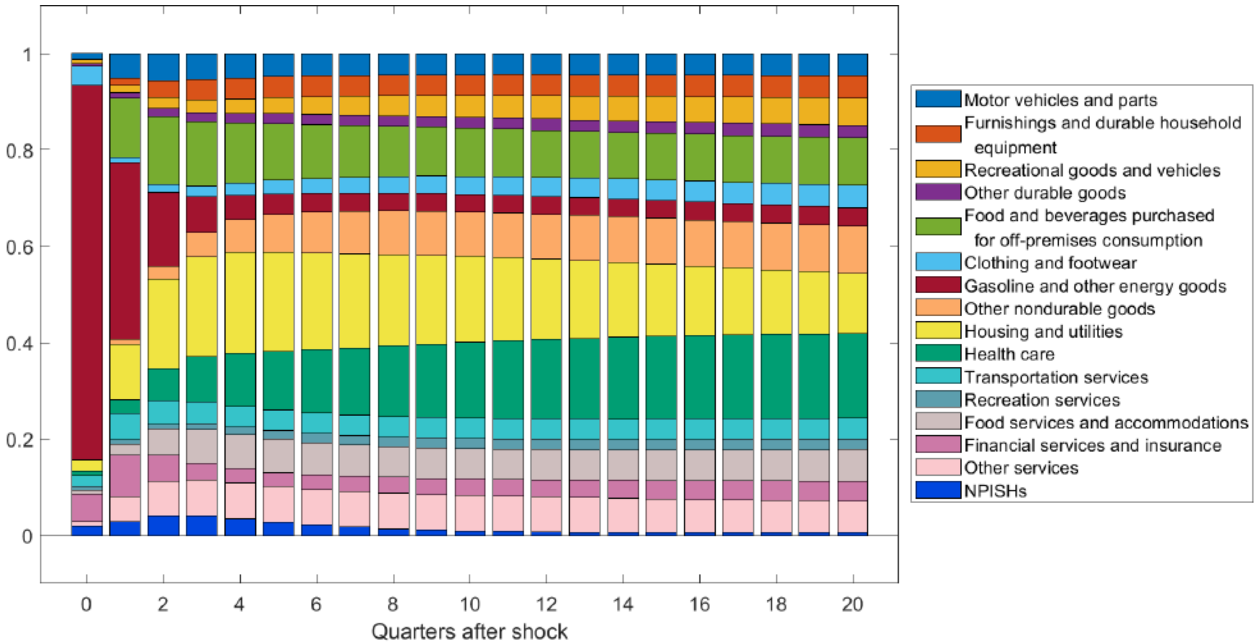

Figure 7: Sectoral Contribution to Aggregate Inflation Persistence

4.4 Sectoral Shocks

4.4.1 Aggregate Impulse Responses to Sectoral Shocks

Similarly, for sectoral shocks, we can also calculate the impulse response of aggregate inflation to each sectoral shock. According to Eq.(5), aggregate IRs to sectoral shocks are only related to the transition matrix $\mathbf{A}_1$. Figure 8 shows the impulse response of aggregate inflation to different sectoral shocks. Surprisingly, there are significant differences in how aggregate inflation reacts to various sectoral shocks. Current period responses to each sectoral shock are naturally equal to sector weights. However, after being impacted by the "Health Care" and "Recreational Goods and Vehicles" sectors, aggregate inflation's response has surprisingly grown for nearly a year. The response on the "Recreational Goods and Vehicles" sector had the highest rise in aggregate inflation IR, rising from a starting point of 0.03% to over a tenth of the shock size. The sectoral shocks in "Housing and Utilities" and "Food Services and Accommodations" also provided a short-lived but significant boost to the reaction of aggregate inflation. The IRs often drop monotonically when off-diagonal elements are disabled, which means sectoral spillover is the cause of the IR curve's rising and then downward tendency.

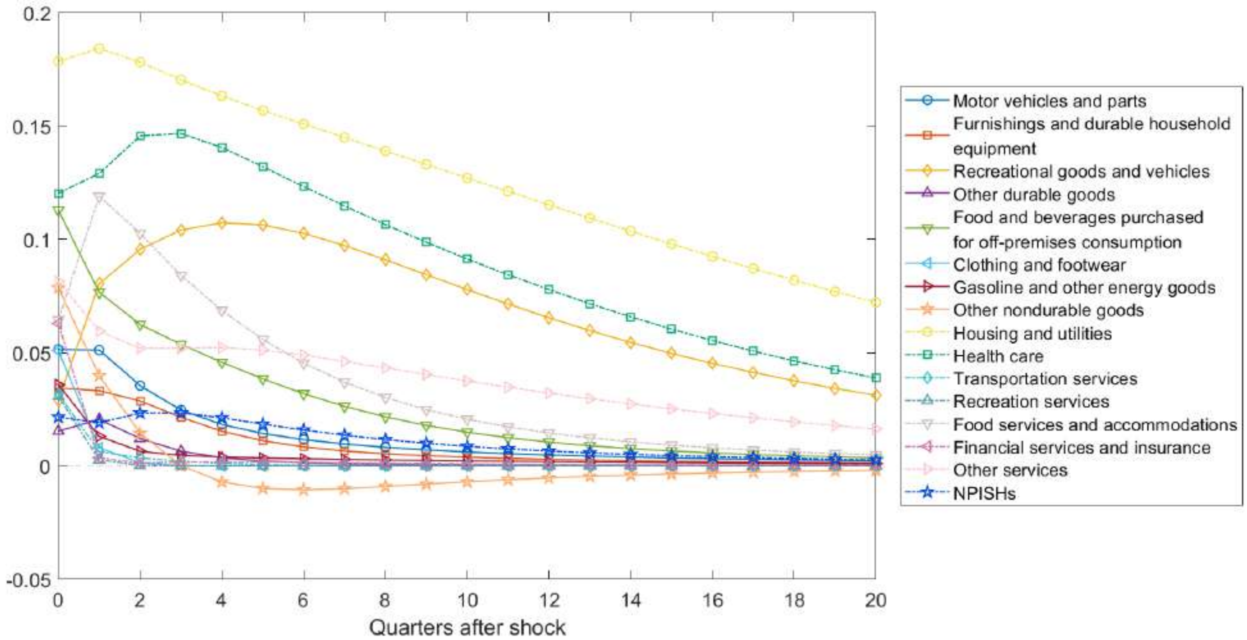

Figure 8: Aggregate Impulse Responses to Sectoral Shocks

4.4.2 Sectoral Impulse Responses to Sectoral Shocks

The sectoral IRs on these four sectors ("Housing and Utilities", "Food Services and Accommodations", "Health Care", and "Recreational Goods and Vehicles"), respectively, are shown in Figures 9-12 to further illustrate how the IR curves form. The examination of the transition matrix in Section 3.3 also reveals the phenomena of these four sectors. The "Housing and Utilities" sectoral shock primarily has a significant and long-lasting effect on its own inflation, remarkably holding half of the shock's size a year later, although it has less of an effect on other sectors. The "Food Services and Accommodations" sector, on the other hand, does not respond well to its own shock, dropping to 0.34% after a quarter. But it instead spreads the shock to a lot of other sectors, like "NPISHs", "Food and Beverages Purchased for Off-premises Consumption", and "Recreation Services", which is the reason why aggregate inflation in Figure 8 can rise rapidly by 0.05% in just one quarter.

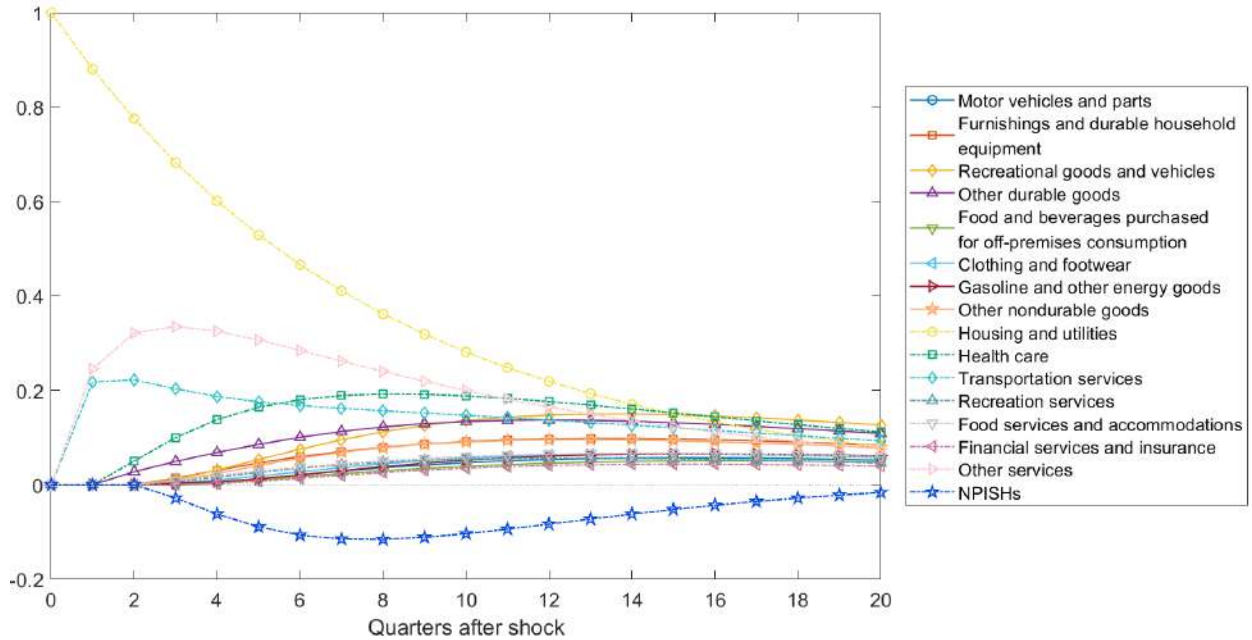

Figure 9: Sectoral Impulse Responses to "Housing and Utilities" Shock

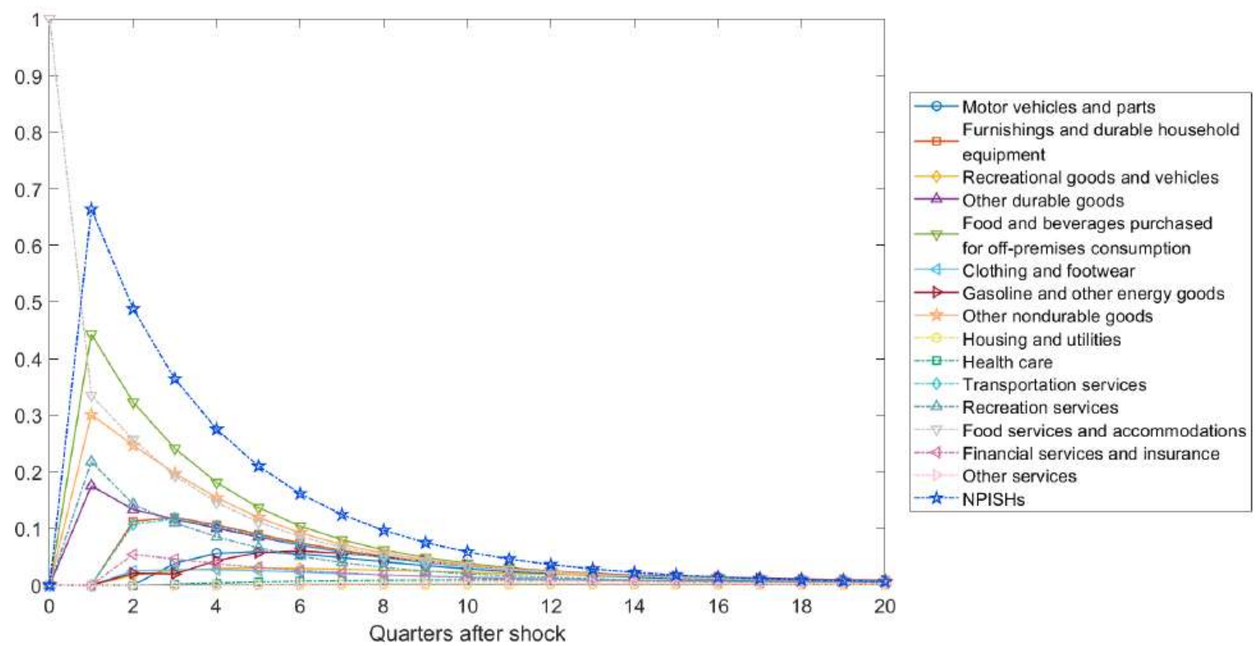

Figure 10: Sectoral Impulse Responses to "Food Services and Accommodations" Shock

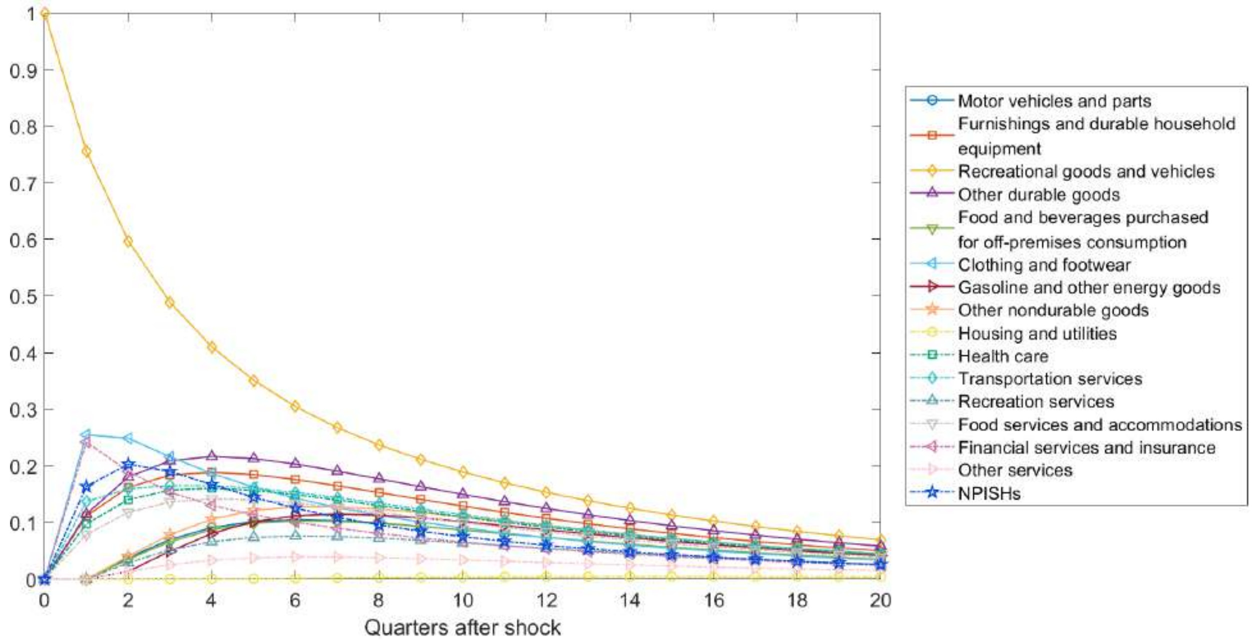

Figure 11: Sectoral Impulse Responses to "Recreational Goods and Vehicles" Shock

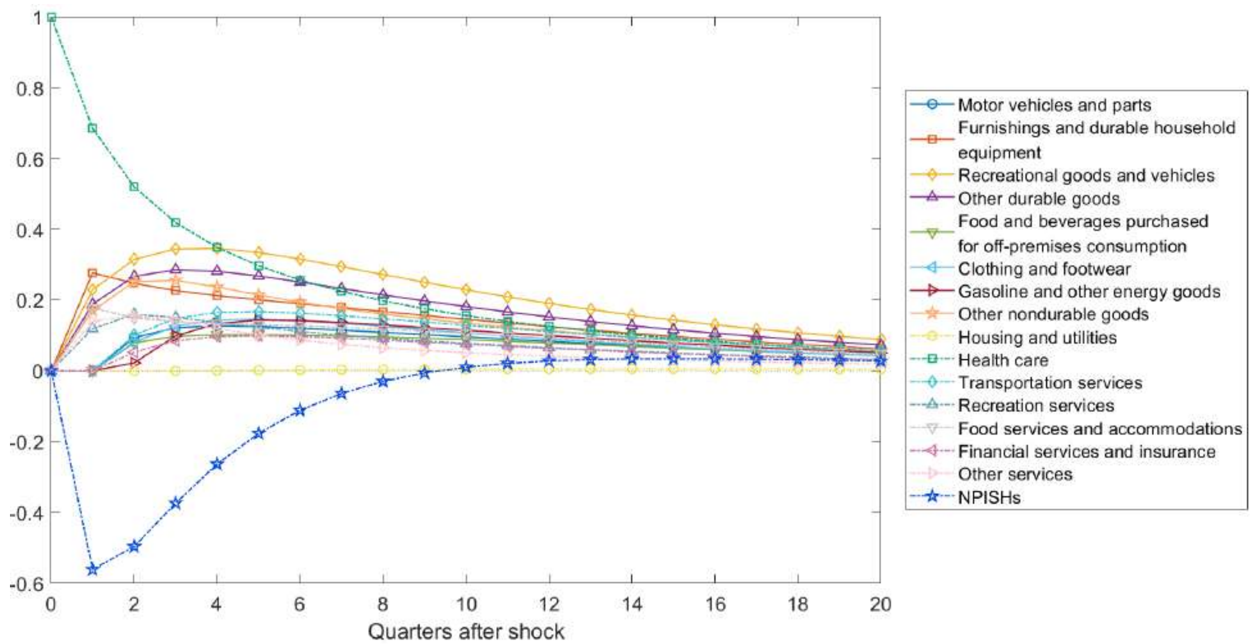

Figure 12: Sectoral Impulse Responses to "Health Care" Shock

The shock of "Health Care" and "Recreational Goods and Vehicles" have relatively similar effects, in which the self-driven and sector spillover channels are parallel. When its own persistence is high, it can also spread persistence to many other sectors, causing linkage effects. Specifically, the sector of "Recreational Goods and Vehicles" affects eight sectors in the first quarter following its shock, including "Furnishings and Durable Household Equipment", "Health Care", and "Food Services and Accommodations". After that, "Motor Vehicles and Parts" after being impacted by "Furnishings and Durable Household Equipment", "Food and Beverages Purchased for Off-premises Consumption" after being impacted by "Food Services and Accommodations", "Recreation Services" after being impacted by "Health Care" and "Food Services and Accommodations", all appear new positive shock response in the second quarter. Consequently, the persistence continued to overflow and spread, culminating in an increase in the aggregate inflation's IR of one year in reaction to the shock of "Recreational Goods and Vehicles", as seen in Figure 8. As seen in Figure 12, spillover effects to seven sectors lead the IR curve to climb in the first quarter despite the negative shock reaction of "NPISHs".

4.5 Decomposition of Aggregate Inflation

4.5.1 Aggregate Inflation and Common Shock Component

We can decompose the aggregate inflation as

$$\pi_t = \boldsymbol{w}^\top \mathbf{Y}_t = \boldsymbol{w}^\top \sum_{s=0}^{\infty} (\mathbf{A}_1)^s (\mathbf{\Lambda f}_{t-s} + \mathbf{u}_{t-s}) = \boldsymbol{w}^\top \sum_{s=0}^{\infty} (\mathbf{A}_1)^s \mathbf{\Lambda f}_{t-s} + \boldsymbol{w}^\top \sum_{s=0}^{\infty} (\mathbf{A}_1)^s \mathbf{u}_{t-s} \quad (12)$$

Using the moving average representation of a VAR model, the aggregate inflation is decomposed into two parts, i.e. the past and current common shock component associated with the common factors and factor loading and the past and current sectoral shock component connected to the sectoral shocks.

Figure 13 shows the demeaned aggregate inflation, π_t versus its common shock component—the first term of the right-most expression of (12). The two components have a relatively strong association (0.6 for 1970Q1-2023Q3), but the former is noticeably more volatile, particularly in times of Great Inflation. This implies that the sectoral shock component continues to contribute to aggregate inflation to some extent. Nevertheless, the impact of the sectoral shock component has been significantly reduced since the common factor was incorporated into the model.

We further decompose the common shock component $\boldsymbol{w}^\top \sum_{s=0}^{\infty} (\mathbf{A}_1)^s \mathbf{\Lambda f}_{t-s}$ into 16 by-sector components, with each in the form of $w_i \left(\sum_{s=0}^{\infty} (\mathbf{A}_1)^s \mathbf{\Lambda f}_{t-s}\right)_{i,*}$, as shown in Figure 14. Obviously, the sector of gasoline and other energy goods makes up the majority of the contributions made to the common shock component, both in the Great Inflation period and in the COVID-19 period. The common factor is closely related to the inflation of crude oil prices (Brent) as shown in Figure 15, with a correlation of 0.41 from 1960Q2 to 2023Q3. The

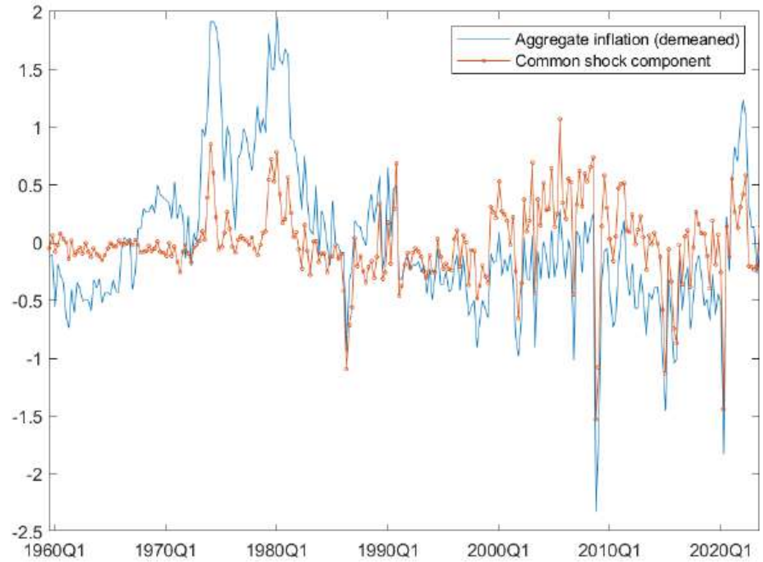

Figure 13: Common Shock Component and Aggregate Inflation (Demeaned)

Notes: Aggregate inflation is demeaned as well because the common factors are estimated from the demeaned sectoral inflation. Before 1970, the calculation of the common shock component is not accurate due to insufficient data.

two have a stronger link during times of high inflation, which is 0.66 and 0.74, respectively, during the Great Inflation period (from 1970Q1 to 1979Q4) and the COVID-19 period (from 2021Q1 to 2023Q3).

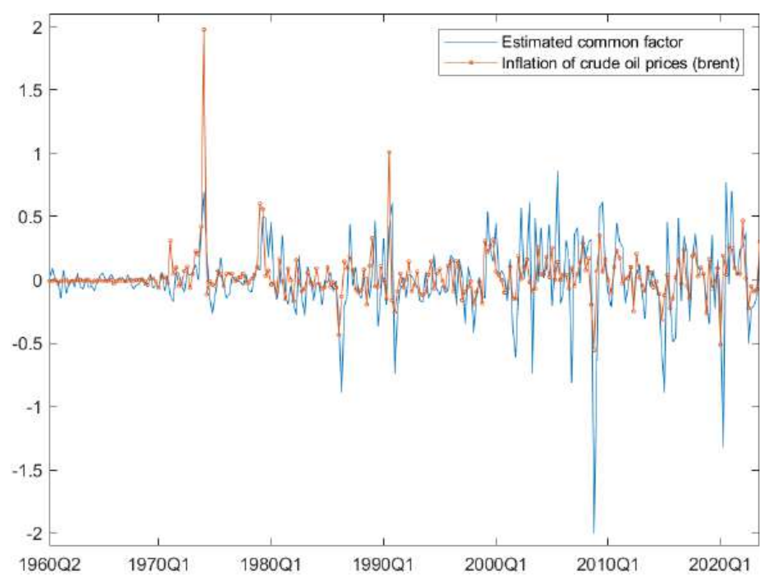

Figure 15: Common Factors and Inflation of Crude Oil Prices (Brent)

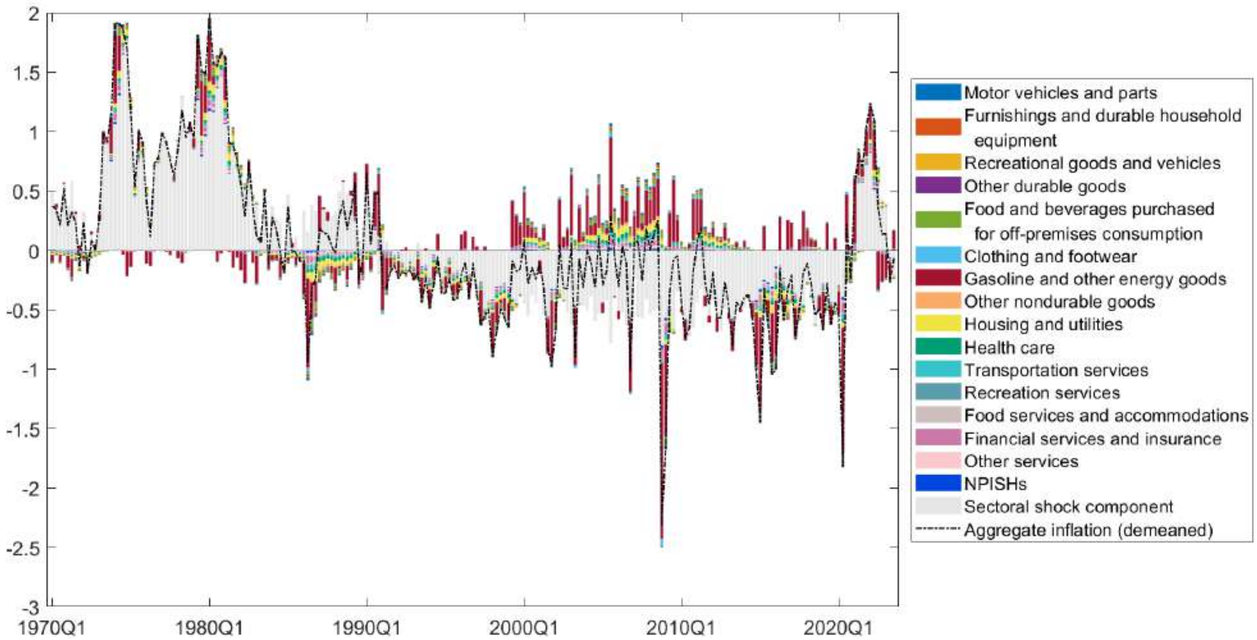

Figure 14: Demeaned Aggregate Inflation, Decomposed Common Shock Components, and the Sectoral Shock Component

4.5.2 Aggregate Inflation with and without spillover

(12) states that the transfer matrix, weights, and error term all affect aggregate inflation. The transfer matrix's off-diagonal components will be zero if sector spillover is not taken into account, creating a diagonal matrix. (13) uses this diagonal matrix to forecast aggregate inflation time series without spillover effects, while maintaining the estimated shock and sectoral weights of the original model.

$$\pi_t^n = \boldsymbol{w}^\top \mathbf{Y}_t^n = \boldsymbol{w}^\top \sum_{s=0}^{\infty} \left(\operatorname{diag}\left(\widehat{\mathbf{A}}_1\right)\right)^s \left(\widehat{\boldsymbol{\Lambda}}\widehat{\mathbf{f}}_{t-s} + \widehat{\mathbf{u}}_{t-s}\right) \tag{13}$$

Figure 16 shows the demeaned aggregate inflation with and without spillover. It is evident that ignoring intersectoral spillover will greatly understate aggregate inflation, particularly during the high inflation period. A diagonal transfer matrix's anticipated inflation value is much lower than the actual inflation rate during the Great Inflation, with the largest discrepancy reaching 1.25% in 1974Q4. The average difference between the value of aggregate inflation without spillover and the actual aggregate inflation rate for 2021Q3-2023Q2 is 0.42%, with the maximum difference of 0.6% happening in 2022Q2.

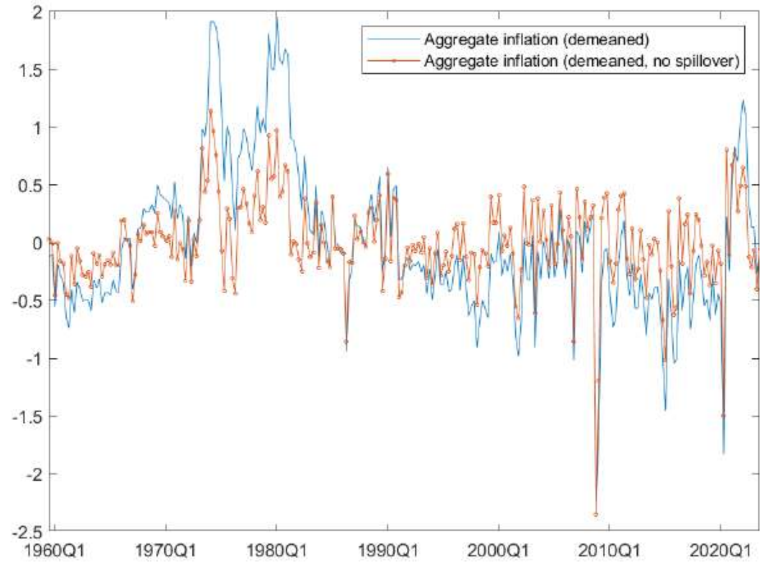

Figure 16: Demeaned Aggregate Inflation with and without spillover

Notes: Aggregate inflation is demeaned because the common factors and sectoral shocks are estimated from the demeaned sectoral inflation.

4.5.3 Sectoral Shock Components in Great Inflation Period

Similarly, we can decompose the sectoral shock component $\boldsymbol{w}^{\top} \sum_{s=0}^{\infty} (\mathbf{A}_1)^s \mathbf{u}_{t-s}$ into each of the 16 by-sector additive components. The decomposition for 1970Q1-1985Q4 is shown in Figure 17.

Figure 17 indicates that the common factor contributes significantly to aggregate inflation during the two periods when aggregate inflation is more than 1%, peaking at 47% in 1979Q3. With respect to the influence of sectoral shocks, "Food and Beverages Purchased for Off-premises Consumption" clearly helps the early period's surge in aggregate inflation, rising from 0.27% in 1972Q1 to 0.99% in a single quarter. Since then, the shock of the "Housing and Utilities" sector has taken center stage, with the "Health Care" and "Financial Services and Insurance" sectors also having a major positive effect.

4.5.4 Sectoral Shock Components in COVID-19 Period

The bars in Figure 18 show the annual sectoral inflation rates and the numbers on the bars represent the annual aggregate inflation equal to the sum of 16 bars. The annual inflation rate of 2023 is obtained by annualizing the inflation rate for three quarters. "Housing and utilities" sector inflation was the primary cause of aggregate inflation in the three years between 2017 and 2019. Prices in the "Food and Beverages Purchased for Off-premises Consumption" sector skyrocketed during COVID-19 in 2020, while sectoral inflation in the "Health Care" also rose. Aggregate inflation did not grow dramatically, though, as a result

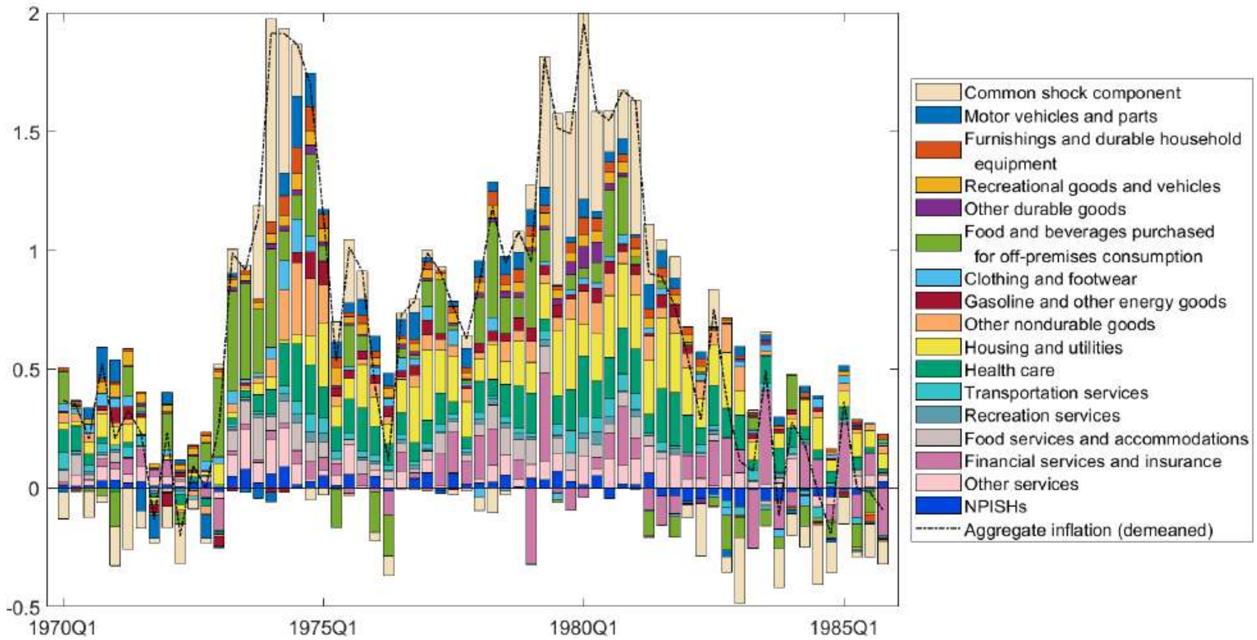

Figure 17: Demeaned Aggregate Inflation, the Common Shock Component, and Decomposed Sectoral Shock Components: Great Inflation Period

of a decline in both global demand and oil prices. International oil prices rose from their lows in 2021 as demand increased and the worldwide pandemic situation improved, pushing inflation to a record high of 6.71% in recent years. Even while inflation in "Gasoline and Other Energy Goods" slowed in 2022, Food and housing costs kept rising, and aggregate inflation remained high until the increase in food prices eased in 2023.

To further explore the reasons for rising inflation, the decomposition of demeaned aggregate inflation for 2020Q1-2023Q3 to the common shock component and 16 by-sector sectoral shock components is shown in Figure 19. The sectoral shock of "Food and Beverages Purchased for Off-premises Consumption" is the first to arise to cause an increase in aggregate inflation in 2020 while the common factor is negative. As a result of the "Food and Beverages Purchased for Off-premises Consumption" sector's spillover impact, other sectors reacted in the third quarter of 2020, which saw a sharp increase in inflation. The sectors affected by the spillover included "Food Services and Accommodations", "Financial Services and Insurance" and "Furnishings and Durable Household Equipment". In 2021, the "Gasoline and Other Energy Goods" sectoral shock would inevitably increase as demand and oil prices rise. According to the transition matrix, inflation in both the "Motor Vehicles and Parts" and "Gasoline and Other Energy Goods" sectors soared significantly as a result of the shock of the former, reaching a peak in aggregate inflation. As the figure illustrates, the primary reason for the high inflation in "Housing and Utilities" and "Food and Beverages Purchased

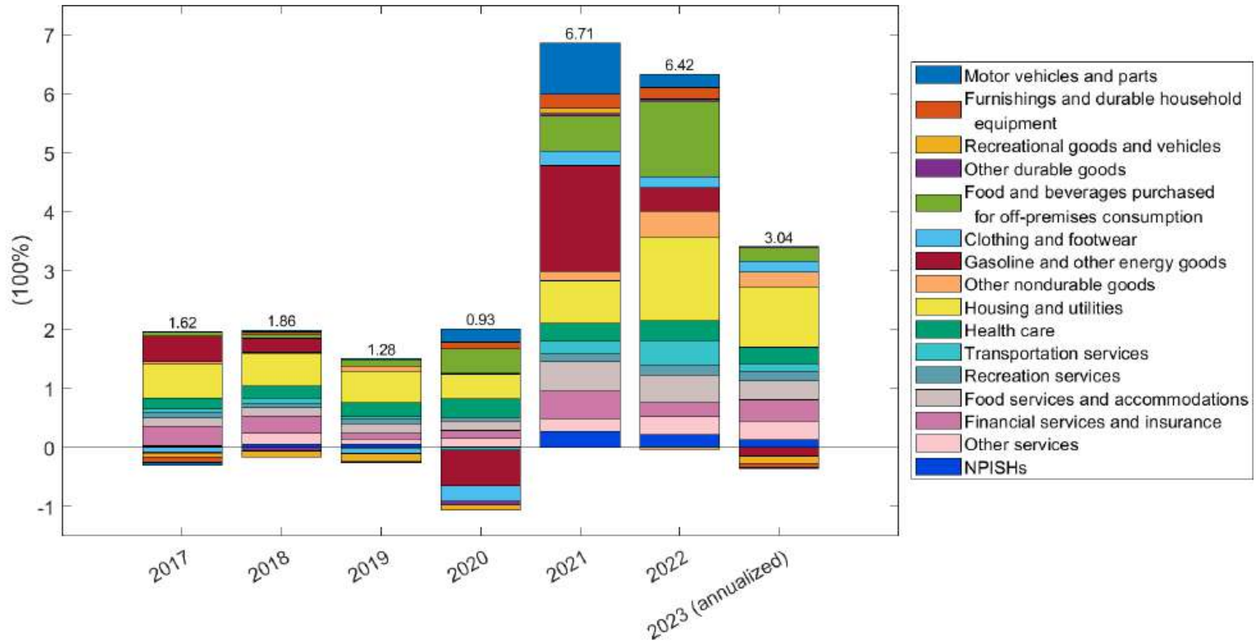

Figure 18: Composition of Aggregate Inflation

Notes: The numbers on the bars represent the annual aggregate inflation equal to the sum of 16 bars. The annual inflation rate of 2023 is obtained by annualizing the inflation rate for three quarters.

for Off-premises Consumption" in 2022 and 2023 is the sectors' own shock.

5 Conclusion

In this paper, we demonstrate that inflation spillover across sectors highlights an important new aspect of inflation persistence. Inflation spillover across sectors is particularly important during high inflation periods. The empirical model we employed to capture inflation spillover is very flexible. The model can be broadly applied to various settings to study the persistence of other macro variables.

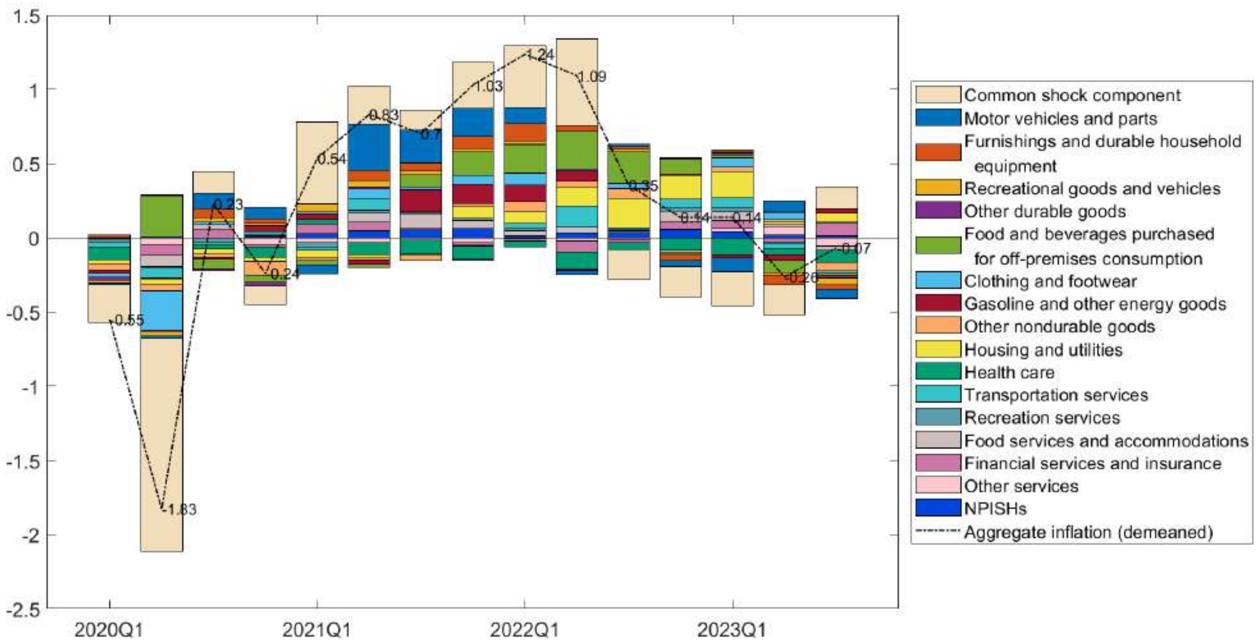

Figure 19: Demeaned Aggregate Inflation, the Common Shock Component, and Decomposed Sectoral Shock Components: COVID-19 Period

References

Acemoglu, D., V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi (2012). The network origins of aggregate fluctuations. *Econometrica 80*(5), 1977–2016.

Afrouzi, H. and S. Bhattarai (2023). Inflation and gdp dynamics in production networks: A sufficient statistics approach. Technical report, National Bureau of Economic Research.

Altissimo, F., B. Mojon, and P. Zaffaroni (2009). Can aggregation explain the persistence of inflation? *Journal of Monetary Economics 56*(2), 231–241.

Baqaee, D. R. and E. Farhi (2019). The macroeconomic impact of microeconomic shocks: Beyond hulten's theorem. *Econometrica 87*(4), 1155–1203.

Benigno, P. and G. B. Eggertsson (2023). Itâs baaack: The surge in inflation in the 2020s and the return of the non-linear phillips curve. Technical report, National Bureau of Economic Research.

Bernanke, B. and O. Blanchard (2023). What caused the us pandemic-era inflation? *Hutchins Center Working Papers*.

Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica 79*(3), 733–772.

Gagliardone, L. and M. Gertler (2023). Oil prices, monetary policy and inflation surges. Technical report, National Bureau of Economic Research.

Guerrieri, V., G. Lorenzoni, L. Straub, and I. Werning (2022). Macroeconomic implications of covid-19: Can negative supply shocks cause demand shortages? *American Economic Review 112*(5), 1437–1474.

Kock, A. B. and L. Callot (2015). Oracle inequalities for high dimensional vector autoregressions. *Journal of Econometrics 186*(2), 325–344.

Lee, J. H., Z. Shi, and Z. Gao (2022). On lasso for predictive regression. *Journal of Econometrics 229*(2), 322–349.

Mayoral, L. (2013). Heterogeneous dynamics, aggregation, and the persistence of economic shocks. *International Economic Review 54*(4), 1295–1307.

Miao, K., P. C. Phillips, and L. Su (2023). High-dimensional vars with common factors. *Journal of Econometrics 233*(1), 155–183.

Minton, R. and B. Wheaton (2023). Delayed inflation in supply chains: Theory and evidence. *Available at SSRN 4470302*.

Rubbo, E. (2023). Networks, phillips curves, and monetary policy. *Econometrica 91*(4), 1417–1455.

Su, L., Z. Shi, and P. C. B. Phillips (2016). Identifying latent structures in panel data. *Econometrica 84*(6), 2215–2264.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology 58*(1), 267–288.

Appendix

The estimation procedure. In our context, we have 256 ($= 16 \times 16$) slope coefficients in the VAR system. Our observation is a panel data of 16 time series with each of the length 257 time periods. We will need regularization to control the quality of estimation.

First-step estimator:

(a) Estimate preliminarily $\mathbf{B}$ and $\mathbf{\Psi}$ by running the following ℓ_1-nuclear-norm regularized regression:

$$\left(\widetilde{\mathbf{B}}, \widetilde{\mathbf{\Psi}}\right) = \underset{(\mathbf{B},\mathbf{\Psi})}{\operatorname{argmin}} \frac{1}{2NT} \left\|\mathbf{Y} - \mathbf{X}\mathbf{B} - \mathbf{\Psi}\right\|_F^2 + \frac{\gamma_1}{N} \left|\operatorname{vec}\left(\mathbf{B}\right)\right|_1 + \frac{\gamma_2}{\sqrt{NT}} \left\|\mathbf{\Psi}\right\|_*$$

where $\mathbf{\Psi} = \mathbf{F}\mathbf{\Lambda}^\top$, γ_1 and γ_2 are tuning parameters. This step gives us consistent estimation of the slope coefficients B when $\gamma_1, \gamma_2 \to 0$ asymptotically. This step will introduce non-trivial shrinkage bias, which produces non-standard distribution in the slope coefficient estimation. In other to estimate the factor and the factor loading, we must also determine the number of factors at the first place. That motivates the second step.

(b) Estimate the number of factors R by singular value thresholding as follows:

$$\widehat{R} = \sum_{i=1}^{\min(N,T)} \mathbb{I}\left\{\psi_i\left(\widetilde{\mathbf{\Psi}}\right) \geq \left(\gamma_2\sqrt{NT}\left\|\widetilde{\mathbf{\Psi}}\right\|_2\right)^{1/2}\right\}$$

where $\psi_i\left(\widetilde{\mathbf{\Psi}}\right)$ are the singular values of $\widetilde{\mathbf{\Psi}}$.

(c) Let the singular value decomposition of $\widetilde{\mathbf{\Psi}}$ be $\widetilde{\mathbf{\Psi}} = \widetilde{\mathbf{U}}\widetilde{\mathbf{D}}\widetilde{\mathbf{V}}^\top$, where $\widetilde{\mathbf{D}} = \operatorname{diag}\left(\psi_1\left(\widetilde{\mathbf{\Psi}}\right), \ldots, \psi_{\min(N,T)}\left(\widetilde{\mathbf{\Psi}}\right)\right)$. Then obtain a preliminary estimate of $\mathbf{F}$:

$$\widetilde{\mathbf{F}} = \sqrt{T}\widetilde{\mathbf{U}}_{*,[\widehat{R}]}$$

The factors is estimated as the left-eigenvectors (standardized by $\sqrt{T}$).

Second-step estimator (Lasso): Since N is of high-dimension, when we estimate the slope coefficients we impose sparsity. Notice that we don't need penalize the factor loadings, which is common in the literature of interactive fixed effect with a large N. For each $i \in N$, solve the Lasso problem:

$$\left(\dot{\mathbf{B}}_{*,i}^{\top}, \dot{\boldsymbol{\lambda}}_{i}^{\top}\right)^{\top} = \underset{\left(\mathbf{B}_{*,i}^{\top}, \boldsymbol{\lambda}_{i}^{\top}\right)^{\top}}{\operatorname{argmin}} \frac{1}{2T} \left\| \mathbf{Y}_{*,i} - \mathbf{X}\mathbf{B}_{*,i} - \widetilde{\mathbf{F}}\boldsymbol{\lambda}_{i} \right\|_{F}^{2} + \gamma_{3} \left| \mathbf{B}_{*,i} \right|_{1}$$

where γ_3 is tuning parameter. (We will need to define $\mathbf{Y}_{*,i}$ as the $(T-1) \times 1$ vector for each individual i.)

Notice the above optimization problem is for a single i (time-fixed effect must be handled prior to this step). There we focus on the consistency, instead of seeking efficiency by exploring the covariance structure of the errors. The equation-by-equation (for each i) estimation is consistent with no problem.

The purpose of the second step is to refresh a preliminary consistent so that in terms of econometrics we can establish selection consistency.

The second-step estimators of $\mathbf{B}$ and $\boldsymbol{\Lambda}$ are given by $\dot{\mathbf{B}} = \left(\dot{\mathbf{B}}_{*,1}, \cdots, \dot{\mathbf{B}}_{*,N}\right)$ and $\dot{\boldsymbol{\Lambda}} = \left(\dot{\lambda}_{1}, \cdots, \dot{\lambda}_{N}\right)^{\top}$.

Third-step estimator (Conservative Lasso):

(a) Set $\widehat{\mathbf{F}}^{(0)} = \widetilde{\mathbf{F}}$. For integer $\ell \geq 1$, update the estimates of $\mathbf{B}$ and $\boldsymbol{\Lambda}$ using:

$$\left(\widehat{\mathbf{B}}_{*,i}^{(\ell)\top}, \widehat{\boldsymbol{\lambda}}_{i}^{(\ell)\top}\right)^{\top} = \underset{\left(\mathbf{B}_{*,i}^{\top}, \boldsymbol{\lambda}_{i}^{\top}\right)^{\top}}{\operatorname{argmin}} \frac{1}{2T} \left\| \mathbf{Y}_{*,i} - \mathbf{X}\mathbf{B}_{*,i} - \widehat{\mathbf{F}}^{(\ell-1)}\boldsymbol{\lambda}_{i} \right\|_{F}^{2} + \gamma_{4} \sum_{k=1, k\neq i}^{N} w_{k,i} \left| B_{k,i} \right|$$

where $w_{k,i} = \mathbf{I}\left(\left|\dot{B}_{k,i}\right| < \gamma_4\right)$, and γ_4 is tuning parameter. (This step is still for each individual i)

(b) Update $\widehat{\mathbf{F}}^{(\ell)}$ the singular value decomposition of $\mathbf{Y} - \mathbf{X}\widehat{\mathbf{B}}^{(\ell)}$. ($F$ is global here so it requires all individuals)

(c) Iterate steps 1-2 for a finite time ℓ^*. Denote the final estimators by $\widehat{\mathbf{B}} = \widehat{\mathbf{B}}^{(\ell^*)}$, $\widehat{\mathbf{F}} = \widehat{\mathbf{F}}^{(\ell^*-1)}$, and $\widehat{\boldsymbol{\Lambda}} = \widehat{\boldsymbol{\Lambda}}^{(\ell^*)}$. (Basically, once F is provided, then estimate B and lambda for each individual separately).