Optimal Monetary Policy in Production Networks with Distortions*

Zhihao Xu† Changhua Yu‡

This version: September 2024

Abstract

This paper studies optimal monetary policy in a multi-sector economy with input-output linkages and distortions. Our model incorporates both supply-side and demand-side effects of monetary policy. We derive a tractable sufficient statistic for the supply-side effect, which comprises two reallocation channels resulting from substitution in production and from substitution in consumption. The optimal monetary policy induces an inflation bias stemming from both an aggregate wedge and the supply-side effect, and stabilizes an inflation index by assigning higher weights to (i) larger sectors, (ii) sectors with stickier prices, and (iii) sectors with less distortions. Our quantitative results indicate that production networks play a crucial role in generating both supply-side and demand-side effects of monetary policy.

Keywords: Production Networks, Optimal Monetary Policy, Supply-Side Effect, Phillips Curves, Stabilization Policy

JEL Codes: E12 E31 E52 E58

*We are grateful to David Baqaee, Mario Crucini, Michael Devereux, Simon Gilchrist, Zhen Huo, Minghao Li, Shaowen Luo, Shengxing Zhang, and seminar and conference participants at various institutions for helpful comments and discussions. We thank Michael Weber for sharing the data. We thank the Editor Guillermo Ordonez, an associate editor and two referees for their insightful and constructive suggestions. Yu acknowledges research support under grant 72173007 from National Natural Science Foundation of China. All errors are our own.

†Unaffiliated. E-mail: zhxu@purdue.edu.

‡China Center for Economic Research, National School of Development, Peking University. Address: 5 Yiheyuan Rd, Haidian, Beijing, China, 100871. E-mail: changhuay@gmail.com.

1 Introduction

The disruptions in supply chains during the COVID-19 pandemic have highlighted the importance of understanding how shocks propagate through production networks and impact the overall economy. It is crucial to recognize that shocks in different industries can have varying effects on aggregate output and inflation. In response to these challenges, monetary authorities in both advanced and emerging economies have implemented aggressive measures to combat inflation. The canonical New Keynesian frameworks primarily focus on the demand-side effects of monetary policy. By tightening monetary policy, aggregate nominal demand is reduced, which leads to stabilization of inflation. However, empirical studies have shown that monetary policy can also influence resource allocation and account for a significant portion of aggregate productivity movements (see for instance, Evans, 1992, Barth and Ramey, 2002, Ravenna and Walsh, 2006, Meier and Reinelt, 2022).

How do demand-side shocks, such as monetary policy shocks, affect an economy's output and productivity in a distorted economy with input-output linkages? What is the optimal conduct of monetary policy under such circumstances? Based on frameworks of Long and Plosser (1983) and Baqaee and Farhi (2020), this paper investigates both the supply-side and demand-side effects of monetary policy in a multi-sector model with nominal rigidities, initial markups, and input-output linkages. We find that the supply-side effect of monetary policy arises from resource reallocation across sectors, and is characterized by a tractable sufficient statistic, which can be further broken down into two reallocation channels: one channel due to substitution in production and the other channel due to substitution in consumption.

The reallocation channel resulting from substitution in production depends on several factors, including the average markup, the covariance between wage pass-throughs to sectoral prices and sectors' upstream markups, elasticities of substitution among inputs, and sectoral Domar weights. As noted in Baqaee et al. (2024), an initial misallocation of resources is a *necessary* condition for the supply-side effect of monetary policy. When the average markup in an economy is higher, there is a greater potential for policy interventions to improve resource allocation. Wage pass-throughs and upstream markups reflect the compound impact of frictions throughout the production chain on sectoral prices and markups, respectively. When sectors with high upstream markups tend to exhibit low wage pass-throughs to prices, and vice versa, sectors with high upstream markups will raise their prices to a less extent than sectors with low upstream markups in response to expansionary monetary policy. Consequently, a reduction in relative prices leads to increased demand

and output in sectors with high upstream markups. As a result, both labor and intermediate inputs will be reallocated from sectors with low upstream markups to sectors with high upstream markups. In a distorted economy, marginal product of input in sectors with high upstream markups is greater than that in sectors with low upstream markups. This reallocation of resources across sectors ultimately contributes to an improvement in total factor productivity.

The elasticity of substitution among inputs also plays a crucial role in the supply-side effects. A higher elasticity of substitution among inputs leads to a more significant reallocation channel. This is because downstream sectors' demand responds more strongly to changes in relative sectoral prices when the elasticity of substitution is greater. As a result, monetary policy exerts a stronger influence on resource reallocation. Additionally, sector size also affects the reallocation channel, and a change in resource allocation in larger sectors contributes more to the whole economy.

The reallocation channel resulting from substitution in consumption operates similarly to that in production, but with a significant distinction due to the role of labor. In consumption, substitution is limited to different sectoral products, whereas in production, labor is a crucial input, and substitution occurs among intermediate inputs and between intermediate inputs and labor. In response to expansionary monetary policy, firms tend to substitute labor for intermediate inputs, as sticky sectoral prices increase less than flexible nominal wages due to incomplete pass-throughs. This reallocation towards intermediate inputs enhances economy-wide allocative efficiency, as these inputs are initially underutilized due to double marginalization. These two reallocation channels contribute to the supply-side effects of monetary policy in specific environments, critically depending on the production networks within an economy. For instance, in a horizontal economy where multiple sectors rely solely on labor as a productive input, the only prevailing channel is reallocation due to consumption substitution. Conversely, in a one-sector roundabout economy, reallocation resulting from substitution in production becomes the sole channel.

In our model economy, labor supply adjusts endogenously to shocks. Consequently, there is a traditional New Keynesian demand-side effect of monetary policy, in addition to the supply-side effect. We provide tractable expressions for both the supply and demand-side effects of monetary policy. The manner in which wages and labor respond to shocks determines these two effects. We demonstrate that the supply-side effect increases with the inverse Frisch elasticity of labor supply, while the demand-side effect decreases with this elasticity. When labor supply is less elastic, a change in nominal wages due to shocks becomes more pronounced, whereas the labor supply itself is less responsive to these shocks. In a model with price rigidities and initial markups, a larger response in nominal wages results in a greater change in ex-post markups, leading to a more significant supply-side effect.

The following section constructs sectoral Phillips curves. The slopes of these curves are determined by input-output linkages, sectoral price rigidities, initial markups, and cross-sector elasticities. Our findings indicate that a positive supply-side effect of monetary policy flattens the slopes of all sectoral Phillips curves. The rationale behind this is that, in response to an increase in the output gap, wages must rise to attract additional labor necessary to support a higher level of production. However, the positive supply-side effects of monetary policy reduce firms' demand for labor, thereby diminishing the required wage increases. This, in turn, lowers input costs and sectoral inflation.

Should the central bank optimally induce an inflation bias to enhance allocative efficiency and increase final output? We further examine the optimal monetary policy in response to sectoral productivity shocks. Up to a second-order approximation of the social welfare function around a flexible-price but distorted steady state, the approximated welfare gains comprise several additive components: a first-order bias resulting from an aggregate wedge and allocative efficiency, and second-order welfare losses stemming from output gap volatility, within-sector and cross-sector price distortions, and variations in allocative efficiency.

In a model economy with multiple sectors and distortionary markups, the central bank's ability to influence the economy is limited by its dependence on a single policy instrument. This instrument must strike a balance between first-order biases and second-order welfare losses. Under optimal monetary policy, the central bank may introduce an inflation bias, which arises from both the supply-side effects of monetary policy and an aggregate wedge, both of which are affected by initial markups.[1] We find that, all else being equal, the optimal monetary policy stabilizes an inflation index by assigning greater weights to larger, stickier, and less distorted sectors. When the supply-side effect is positive, the central bank has an incentive to enhance allocative efficiency by increasing sectoral inflation. Additionally, due to the presence of an aggregate wedge, the central bank may also raise inflation to bring output closer to its efficient level.

We then apply our theoretical framework to data, and quantitatively explore the optimal monetary policy. To achieve this, we utilize the input-output tables provided by the Bureau of Economic Analysis in the USA and map them into our model to obtain a cost-based input-

[1] The initial state of the economy is inefficient due to lacking of enough tax instruments that can fix sectoral monopolistic markups (see for instance, Adão et al. (2003)). The monetary authority faces a trade-off between stabilizing inflation (second-order welfare losses) and substituting for these missing tax instruments (first-order biases). Note also that initial markups are necessary to generate both the supply-side effect and the aggregate wedge in our model economy.

output matrix, labor input shares, and consumption shares from 1997 to 2015. The initial markups at the industry level are taken from Baqaee and Farhi (2020), while the sectoral price rigidities are derived from Pasten et al. (2020). Additionally, industry-level productivities are obtained from the Integrated Industry-Level Production Account of the U.S. Bureau of Economic Analysis and Bureau of Labor Statistics.

Our quantitative analysis demonstrates that monetary policy and production networks play crucial roles in shaping the supply side of the economy. A one-percentage-point increase in nominal wages resulting from expansionary monetary policy leads to an average rise of 0.018% in total productivity. The inflation bias stemming from the supply-side effect moves closely with that arising from the aggregate wedge, varying over time from 0.08% during years of economic downturns to 0.54% during years of economic booms, with an average inflation bias of 0.30%. The welfare analysis illustrates that the supply-side effect generates a similar welfare gain as that induced by the aggregate wedge. When we completely eliminate input-output linkages, the supply-side effect is solely determined by the reallocation resulting from substitution in consumption. The quantitative results show that the supply-side effect and inflation bias associated with the supply-side effect decline substantially in an economy without input-output linkages. Inflation bias due to the aggregate wedge decreases, but it significantly dominates the inflation induced by the supply-side effect.

Can the optimal monetary policy be implemented by a simple monetary policy rule? We examine two alternative simple rules: an output gap targeting rule, as in La'O and Tahbaz-Salehi (2022) and Rubbo (2023), and a CPI inflation targeting rule. Our quantitative analysis reveals that while both output gap stabilization and CPI inflation targeting policies incur less second-order welfare losses, they still underperform compared to optimal monetary policy due to their inability to fully leverage reductions in the aggregate wedge and the benefits of allocative efficiency. Furthermore, the output gap stabilization rule yields higher welfare than the CPI inflation targeting rule. This is because the output gap targeting rule partially accounts for changes in output resulting from allocative efficiency, and stabilizing this output gap leads to lower welfare losses compared to the CPI inflation targeting rule.

Related Literature. Our paper is part of the growing literature that investigates production networks in macroeconomics (Long and Plosser, 1983; Basu, 1995; Acemoglu et al., 2012).[2] The framework of this paper builds on Baqaee and Farhi (2020), who studied the

[2]The pioneering work includes Jones (2011), Gabaix (2011) and Acemoglu et al. (2015). More recent studies include Acemoglu et al. (2017), Atalay (2017), Baqaee (2018), Bouakez et al. (2018), Baqaee and Farhi (2019, 2022, 2024), Levchenko et al. (2019), Liu (2019), Acemoglu and Azar (2020), Acemoglu and Tahbaz-Salehi (2020, 2024), Bigio and La'O (2020), Flynn et al. (2020), Luo (2020), Carvalho et al. (2021), Miranda-Pinto and Young (2022),

implications of exogenous markups in economies with input-output linkages. Compared with Baqaee and Farhi (2020), this paper examines both the supply-side and demand-side effects of monetary policy and characterizes optimal monetary policy in a multi-sector general equilibrium economy with nominal rigidities and distortionary markups.

In previous work on monetary policy and production networks, researchers have focused on the demand side of monetary policy in multi-sector New Keynesian economies (see, for instance, Galí, 2015; Nakamura and Steinsson, 2010; Carvalho and Nechio, 2011; Pasten et al., 2017, 2020; Ghassibe, 2021a).[3] Several recent papers have explored optimal monetary policy in economies with production networks. La'O and Tahbaz-Salehi (2022) characterized the optimal policy in terms of an economy's production network and the extent and nature of nominal rigidities. In a parallel study, Rubbo (2023) emphasized that introducing intermediate inputs reduces the slope of all sectoral and aggregate Phillips curves in a dynamic multi-sector model and derived a novel divine coincidence index that the central bank should target.[4] Along this line of research, Afrouzi and Bhattarai (2023) explored how production linkages amplify the persistence of inflation and GDP responses in multi-sector dynamic models. Our model builds on similar multi-sector New Keynesian economies, but takes into account of sectoral initial markups. These initial markups and production networks help generate the supply-side effect of monetary policy.

This paper also contributes to the literature on the supply-side effect of monetary policy. The literature has documented evidence of monetary policy on the supply-side of economies. Evans (1992) found that monetary and fiscal policies Granger-cause measured Solow residuals, and aggregate demand contributes between one quarter and one half to the variance of these residuals. Barth and Ramey (2002) presented evidence that monetary policy affects the cost of production and consequently aggregate productivity. Meier and Reinelt (2022) documented that monetary policy shocks increase markup dispersion across firms and firms with stickier prices have higher markups. On the theoretical side, Ravenna and Walsh (2006) presented a model with a cost channel for monetary policy. David and Zeke (2021) studied business cycle dynamics in a heterogeneous firm economy, and found that resource allocation can strengthen counter-cyclical monetary policy. Meier and Reinelt (2022) showed

Devereux et al. (2023), Osotimehin and Popov (2023), and Pellet and Tahbaz-Salehi (2023). Also see the recent surveys by Carvalho and Tahbaz-Salehi (2019) and Baqaee and Rubbo (2023).

[3]See more recent studies by Altinoglu (2021), Ghassibe (2021b), Giovanni and Hale (2022), and Luo and Villar (2023).

[4]These studies found that optimal monetary policy isn't able to implement the efficient flexible price allocation even if sector specific tax instruments are used to eliminate distortionary markups at the flexible price/wage equilibrium, which starkly contrasts with the canonical one-sector New Keynesian model (see for instance, Correia et al. (2008), Angeletos and La'O (2020)).

that monetary policy shocks can generate substantial fluctuations in aggregate productivity. Our study complements the existing literature by providing a theoretical exploration of the supply-side effect in a multi-sector New Keynesian economy. A recent related paper to our study is Baqaee, Farhi and Sangani (2024), who studied an economy with heterogeneous firms and endogenous markups and found that monetary policy has a first-order effect on aggregate productivity by using Kimball (1995) preference. Our work differs from theirs in two main aspects. First, this paper focuses on the role of production networks in generating the supply-side effect of monetary policy. Even if all firms have identical initial markups, there may still be heterogeneity in upstream distortions due to double marginalization along the supply chain. Second, our paper explores the optimal monetary policy in an economy with production networks.

Structure of the paper. Section 2 introduces a baseline multi-sector model with input-output linkages, and explores cost pass-throughs through production networks. Section 3 provides a tractable sufficient statistic to capture the supply-side effect of monetary policy. Section 4 shows that the supply-side effect of monetary policy flattens the slope of sectoral Phillips curves. Section 5 derives the optimal conduct of monetary policy. Quantitative results are reported in section 6. Section 7 concludes. All proofs are delegated to the Online Appendix.

2 Model

We start with a static model with N industries, and a primary factor, labor. In each industry $i \in \mathcal{N} \equiv \{1, 2, ..., N\}$, there is a unit mass of monopolistically competitive firms. These firms hire labor and use N intermediate inputs to produce goods, which can either be used as intermediate inputs or consumed by households. A representative household supplies labor to firms and consumes a basket of sectoral goods produced by different industries. Nominal price rigidity restricts firms' ability to fully adjust their prices in response to shocks. The monetary authority controls the supply of money. This model allows us to analyze both the supply-side effect and the traditional demand-side of monetary policy in a New Keynesian economy. Armed with this model, we further explore how production networks facilitate the transmission of monetary policy and shape the optimal conduct of monetary policy.

2.1 Households

The utility function that a representative household maximizes is as follows,

$$U(Y, L) = \frac{Y^{1-\gamma}}{1-\gamma} - \frac{L^{1+\varphi}}{1+\varphi},$$

where γ and φ denote the relative risk aversion and the inverse Frisch elasticity of labor supply, respectively. L is the labor supply and Y is the aggregate final output, which is determined by a constant-returns-to-scale aggregator of final goods produced by N industries,

$$Y = C(c_1, ..., c_N) \tag{1}$$

where c_i represents the quantity of sectoral product i consumed by households, and C denotes a consumption aggregator with a constant elasticity of substitution σ_C.

The representative household's budget constraint is given by

$$\sum_{i=1}^{N} p_i c_i \leq wL + \sum_{i=1}^{N} \Pi_i - T,$$

where p_i and c_i are price and quantity of good i, w denotes nominal wages, Π_i is the profit form sector i, and T is a lump-sum tax collected by the government. The left-hand side of the inequality above is equal to the household's nominal expenditure, and the right-hand side shows various sources of the household's nominal income, including labor income, profits from owning firms, and a lump-sum transfer.

2.2 Firms

There exist a unit mass of firms indexed by $\nu \in [0, 1]$ in each sector. A competitive sectoral bundler aggregates varieties produced by all the firms within sector i into a sectoral output using a constant elasticity of substitution (CES) aggregator,

$$y_i = \left(\int_0^1 y_{i,\nu}^{\frac{\varepsilon_i - 1}{\varepsilon_i}} d\nu \right)^{\frac{\varepsilon_i}{\varepsilon_i - 1}},$$

where the within-sector elasticity of substitution is $\varepsilon_i > 1$. The optimal demand for each variety is $y_{i,\nu} = (p_{i,\nu}/p_i)^{-\varepsilon_i} y_i$, and sectoral price reads,

$$p_i = \left(\int_0^1 p_{i,\nu}^{1-\varepsilon_i} d\nu \right)^{\frac{1}{1-\varepsilon_i}}. \tag{2}$$

For simplicity, we assume that all firms within each sector have the same constant returns to scale (CRS) production technology, which uses labor $L_{i,\nu}$ and intermediate inputs $\{x_{ij,\nu}\}_{j=1}^N$ to produce output $y_{i,\nu}$,

$$y_{i,\nu} = A_i F_i(L_{i,\nu}, \{x_{ij,\nu}\}_{j=1}^N),$$

where A_i is industry-specific productivity. Since all firms within an industry use the same CRS production technology, the cost-minimization problems faced by these firms are identical and therefore the marginal cost can be written as,

$$mc_i \equiv \min_{\{x_{ij}\}_{j=1}^N, L_i} wL_i + \sum_{j=1}^N p_j x_{ij},$$

subjective to a unit production constraint $A_i F_i(L_i, \{x_{ij}\}_{j=1}^N) = 1$. When a firm has a chance to reset its price, the optimal price and corresponding effective markup are given by, [5]

$$p_i = \bar{\mu}_i mc_i \quad \text{and} \quad \bar{\mu}_i = (1-\tau_i)\frac{\varepsilon_i}{\varepsilon_i - 1},$$

where τ_i is a rate of proportional input subsidies (taxes) that the government pays to firms in industry i. Markup distortions may exist if policymakers lack a complete set of industry-specific tax instruments to fully offset monopolistic markups.

2.3 Nominal Rigidities

The literature, for instance, Bils and Klenow (2004), Nakamura and Steinsson (2008, 2010) and Pasten et al. (2020), shows that there exists notable variation in price change frequencies across different industries or goods. To capture industry-specific heterogeneity in nominal price resetting, we allow for firms in different sectors to adjust their prices with different frequencies. For simplicity but without loss of generality, we assume that firms in sector i are randomly and mutually independently drawn with probability $\theta_i \in [0, 1]$ to freely reset their

[5] The optimal price is given by: $p_i^* = \text{argmax}_{p_{i,\nu}} \left[(p_{i,\nu} - (1-\tau_i)mc_i) y_i \left(\frac{p_{i,\nu}}{p_i}\right)^{-\epsilon_i} \right]$.

product prices, while the remaining $1-\theta_i$ fraction of firms will keep their prices unchanged. Firms who can adjust their nominal prices will optimally set markup $\bar{\mu}_i$ over marginal cost. Within each sector, the sectoral nominal price is an aggregate of the reset and unchanged prices, which only partially responds to a change in sectoral marginal cost up to a first order log-linearization,

$$\mathrm{d}\log p_i = \int \mathrm{d}\log p_{i,\nu} d\nu = \theta_i \mathrm{d}\log mc_i.$$

In sectors with high price rigidity (low θ_i), sectoral prices are less responsive to changes in their marginal cost.[6] The realized markup in sector i then becomes,

$$\mathrm{d}\log\mu_i \equiv \mathrm{d}\log(p_i/mc_i) = -(1/\theta_i - 1)\mathrm{d}\log p_i = -(1-\theta_i)\mathrm{d}\log mc_i. \quad (3)$$

Due to nominal rigidities ($\theta_i \leq 1$), an increase in marginal cost leads to a decrease in realized markup.

2.4 Policy Instruments

The policy maker may have access to a set of fiscal policy instruments, for instance, industry-specific subsidies (or taxes), to offset monopolistic markups. We assume that the government funds its subsidies to production by levying a lump-sum tax T on households, maintaining a balanced budget,

$$T = \sum_{i=1}^{N} \frac{\tau_i}{1-\tau_i} p_i y_i.$$

Nevertheless, when there are not enough policy instruments to fully offset monopolistic markups, the allocation in the economy is inefficient even in the absence of nominal price rigidities. In the following analysis, we take a more practical view and focus on the scenario in which policy maker isn't able to fully offset sectoral monopolistic markups. Therefore, the initial allocation of the economy is distorted.

The monetary authority chooses money supply M, which in turn affects nominal prices and wages via a cash-in-advance constraint as in Pasten et al. (2020), La'O and Tahbaz-Salehi (2022) and Devereux et al. (2023):

$$\sum_{i=1}^{N} p_i c_i = P^Y Y = M,$$

[6] Our price resetting mechanism mirrors the sticky price model as outlined in Rubbo (2023) for static scenarios, and aligns with the sticky information approach described in La'O and Tahbaz-Salehi (2022).

where P^Y is the consumer price index defined as $\min_{\{c_i\}_{i=1}^N} \{\sum_{i=1}^N p_i c_i : Y = 1\}$. Note that $P^Y Y$ represents the total consumption expenditure and also nominal final output (GDP).

2.5 Equilibrium

In our baseline model, labor is able to move across sectors without any restrictions. Therefore, the market clearing condition for labor is given by

$$L = \sum_{i=1}^{N} L_i = \sum_{i=1}^{N} \int_0^1 L_{i,v} dv.$$

The output in sector i is either used by firms as inputs for production or consumed by households,

$$y_i = c_i + \sum_{j=1}^{N} x_{ji} = c_i + \sum_{j=1}^{N} \int_0^1 x_{ji,v} dv,$$

for all $i \in \mathcal{N}$.

The equilibrium is defined as a set of variables, including total output (Y), labor supply (L), sectoral outputs ($\{y_i\}_{i=1}^N$), intermediate inputs ($\{x_{ij}\}_{i,j=1}^N$), labor demands ($\{L_i\}_{i=1}^N$), final consumption ($\{c_i\}_{i=1}^N$), consumer price (P^Y), sectoral prices ($\{p_i\}_{i=1}^N$) and nominal wages (w), given exogenous price adjustment probabilities ($\{\theta_i\}_{i=1}^N$), initial markups ($\{\bar{\mu}_i\}_{i=1}^N$) determined by industry-specific taxes/subsidies ($\{\tau_i\}_{i=1}^N$) and within-sector substitution elasticities ($\{\epsilon_i\}_{i=1}^N$), and productivities ($\{A_i\}_{i=1}^N$), such that: (i) in each sector, firms optimally choose intermediate inputs and labor demand to minimize their costs, and optimally reset their prices when they have a chance to adjust; (ii) consumers optimally choose consumption and supply labor given sectoral prices and wages; (iii) the government chooses fiscal instruments $\{\tau_i\}_{i=1}^N$, and the monetary authority sets monetary supply M; (iv) all markets clear.

An increase in monetary supply M drives up nominal wages and consequently increases nominal marginal cost of all firms. From the perspective of firms, their marginal cost critically depends on labor cost since labor is the only primary input in the model economy. Following Baqaee et al. (2024), we treat nominal wages w as the monetary policy instrument instead to simplify expressions.[7] An increase in nominal wages corresponds to an expansionary monetary policy and vice versa.

[7]The monetary authority can equivalently choose either nominal wages or money supply as its instrument. Lemma 6 in the Appendix, shows that there is an isomorphism between setting nominal wages and money supply.

2.6 Production Network and Cost Pass-through

In this section, we examine how production networks and nominal price rigidities alter the transmission of labor and intermediate input costs. We approximate the model around a steady state with initial markups $\{\bar{\mu}_i\}_{i=1}^N$, and obtain a set of log-linearized equations. We then define three types of input-output matrices that will be used throughout our analysis. The first two types are cost-based and revenue-based input-output matrices, which are the same as those in Baqaee and Farhi (2020). The third type, called rigidity-adjusted input-output matrix, closely follows La'O and Tahbaz-Salehi (2022) and Rubbo (2023). Table 1 presents these input-output matrices, their associated Leontief inverse matrices, and Domar weights.

Table 1: Input-Output Matrices

	Cost-based	Revenue-based	Rigidity-adjusted
Consumption share	$b_i = \frac{p_i c_i}{\sum_j p_j c_j}$	-	-
Input-output matrix	$\tilde{\Omega}_{ij} = \frac{p_j x_{ij}}{mc_i y_i}$	$\Omega_{ij} = \bar{\mu}_i^{-1}\tilde{\Omega}_{ij}$	$\hat{\Omega}_{ij} = \theta_i \tilde{\Omega}_{ij}$
Leontief inverse	$\tilde{\Psi} = (I - \tilde{\Omega})^{-1}$	$\Psi = (I - \Omega)^{-1}$	$\hat{\Psi} = (I - \hat{\Omega})^{-1}$
Domar weight	$\tilde{\lambda}' = b'\tilde{\Psi}$	$\lambda' = b'\Psi$	$\hat{\lambda}' = b'\hat{\Psi}$

Notes: Definitions of various input-output matrices and their associated Leontief inverse matrices. To make the notation compact, we treat labor as an additional producer, who sells aggregate labor to producers of products. The associated parameters for the labor sector are $b_L = 0$ and $\bar{\mu}_L = \theta_L = 1$. We then construct $(N+1)\times(N+1)$ input-output matrices where the first N rows and columns correspond to goods, while the last row and column correspond to labor. Accordingly, p_{N+1} and w are used interchangeably to denote wages, and $x_{i(N+1)}$ or L_i to represent labor used by sector i. In addition, $\tilde{\Lambda}_L$, Λ_L, and $\hat{\Lambda}_L$ stand for the Domar weight of labor associated with three input-output matrices respectively.

Cost-based input-output matrix. The entry in the cost-based input-output matrix $\tilde{\Omega}_{ij}$, encodes the cost share of input produced by sector j (or labor input when $j = N+1$) in total cost of sector i. This share also measures the elasticity of marginal cost of producers in sector i with respect to output price of sector j by the Shephard's lemma. Its associated Leontief inverse $\tilde{\Psi} = (I - \tilde{\Omega})^{-1} = I + \tilde{\Omega} + \tilde{\Omega}^2 + \tilde{\Omega}^3 + \cdots$ captures the aggregate elasticity of marginal cost in sector i with respect to output price in sector j through direct and indirect use of intermediate inputs in the production network. The last column of $\tilde{\Omega}$, denoted by $\tilde{\Omega}_{(L)}$, measures the elasticity of sector i's marginal cost with respect to wages. Similarly the aggregate elasticity of marginal cost in sector i with respect to wages is captured by the i-th row of last column of the $\tilde{\Psi}$, denoted as $\tilde{\Psi}_{iL}$.

The elasticity of consumer price P^Y with respect to output price in sector i is captured by consumption share b_i. Then the cost-based Domar weight $\tilde{\lambda}_i$ can be interpreted as a consumption-share weighted aggregate elasticity of marginal cost with respect to sectoral price p_i (or wages w when $i = N+1$). When the resource allocation in the economy is efficient (without markups and nominal rigidities) and there exists only one primary input, labor, the value added in each sector then is solely contributed by labor income, and the elasticity of consumer price P^Y with respect to wages then can be written as $\mathrm{d}\log P^Y/\mathrm{d}\log w = b'\tilde{\Psi}_{(L)} \equiv \tilde{\Lambda}_L = 1$ since $\tilde{\Psi}_{iL} = 1$ for all sectors.

Revenue-based input-output matrix. The element of revenue-based input-output matrix Ω_{ij}, represents the share of sector i's expenditure on input produced by sector j to sector i's total revenue. Let $\bar{\mu}^{-1}$ be a diagonal matrix whose ii-th diagonal element is $1/\bar{\mu}_i$. Similar to the cost-based Leontief matrix, the revenue-based Leontief matrix $\Psi = I + \bar{\mu}^{-1}\tilde{\Omega} + (\bar{\mu}^{-1}\tilde{\Omega})^2 + (\bar{\mu}^{-1}\tilde{\Omega})^3 + \cdots$ captures both direct and indirect use of intermediate inputs through the production network. In particular, the last column of the Leontief inverse, denoted by $\Psi_{(L)}$, records the total payments to labor as a share of sales in each sector by taking into account the fact that intermediate inputs are produced by labor as well. Similarly, the revenue-based Domar weight λ_j reflects the total exposure of households to sector j by taking into account of consumption expenditure shares and input-output linkages. This Domar weight also coincides with the sales share for each sector. The revenue-based Domar weight of labor Λ_L, precisely gives the labor income share in nominal output, $\Lambda_L = \frac{wL}{\sum_i p_i c_i}$.

When evaluated at an inefficient equilibrium (i.e., $\bar{\mu} \geq 1$), both Ψ_{iL} and Λ_L are bounded between 0 and 1. Moreover, they weakly decrease in sectoral initial markups $\{\bar{\mu}_i\}_{i=1}^N$: a higher degree of distortion in the supply chain of sector i leads to a lower value of Ψ_{iL} and Λ_L. Specifically, Ψ_{iL} captures the faction of accumulated direct and indirect labor cost along the production chain in sectoral sales, and therefore, Ψ_{iL} can serve as a measure of upstream markup faced by sector i (lower Ψ_{iL} implies a higher upstream markup). Note also that for each sector, the upstream markup Ψ_{iL} is no larger than its counterpart in an economy without input-output linkages, $\bar{\mu}_i^{-1}$, implying that production networks amplify upstream markups through double marginalization.[8] Labor income share Λ_L reflects an average distortion in the economy since labor is the only primary input. The higher distortion an economy experiences, the lower labor income share.

Rigidity-adjusted input-output matrix. Nominal price rigidities limit nominal price responses to cost changes. To better understand how input costs are transmitted through a production chain in a network economy with these rigidities, we adopt the methodologies proposed by La'O and Tahbaz-Salehi (2022) and Rubbo (2023) by introducing a new input-

[8]See Lemma 5 in Online Appendix.

output matrix, the 'rigidity-adjusted input-output matrix' $\hat{\Omega}$. Each element $\hat{\Omega}_{ij}$ of this matrix quantifies the direct pass-through of sector j's price change to sector i's price change,

$$\hat{\Omega}_{ij} \equiv \frac{\partial \log p_i}{\partial \log p_j} = \frac{\partial \log p_i}{\partial \log mc_i}\frac{\partial \log mc_i}{\partial \log p_j} = \theta_i \tilde{\Omega}_{ij}.$$

The associated Leontief inverse matrix $\hat{\Psi} = I + \Theta\tilde{\Omega} + (\Theta\tilde{\Omega})^2 + (\Theta\tilde{\Omega})^3 + \cdots$ measures the overall pass-through of one sector's price to another sector's price, directly and indirectly, through the production network. Analogously, the direct pass-throughs of nominal wages to prices are captured by the last column of rigidity-adjusted input-output matrix, denoted by $\hat{\Omega}_{(L)}$, and the overall wage pass-throughs are given by the corresponding Leontief inverse $\hat{\Psi}_{(L)}$, which accounts for the total exposure of sectors to labor cost. Finally, by using consumption shares as weights, rigidity-adjusted Domar weight of labor $\hat{\Lambda}_L$ aggregates the overall wage pass-through $\hat{\Psi}_{(L)}$ across sectors, and also reflects the overall wage pass-through into consumer price since $\mathrm{d}\log P^Y/\mathrm{d}\log w = \hat{\Lambda}_L$.

Similar to the revenue-based input-output matrix, sectoral wage pass-through $\hat{\Psi}_{iL}$ is no larger than its own probability of price adjustment θ_i, for all i. The use of intermediate inputs from sticky upstream sectors results in more sluggish nominal price adjustments (Basu, 1995). The pass-through of nominal wages into sectoral price $\hat{\Psi}_{iL}$ and consumer price $\hat{\Lambda}_L$ weakly increases in price adjustment probability $\{\theta_i\}_{i=1}^N$, with both measures bounded between 0 and 1.

Additional notations. We introduce additional notations to simplify our analysis. The superscript n denotes a specific segment within a matrix or a vector. For instance, Ψ^n denotes an $N \times N$ sub-matrix consisting of the first N rows and columns of matrix Ψ. Similarly, $\Psi^n_{(L)}$ refers to an $N \times 1$ vector containing the first N elements of $\Psi_{(L)}$. $\mathrm{d}\log p$ is an $(N+1) \times 1$ vector, with last component $\mathrm{d}\log p_{N+1} = \mathrm{d}\log w$. π denotes sectoral price inflation (the first N components of vector $\mathrm{d}\log p$). We introduce the following weighted covariance operator similarly to Baqaee and Farhi (2020),

$$\mathrm{Cov}_{\tilde{\Omega}(j,:)}\left(\hat{\Psi}_{iL}, \Psi_{iL}\right) = \sum_i \tilde{\Omega}_{ji}\hat{\Psi}_{iL}\Psi_{iL} - \left(\sum_i \tilde{\Omega}_{ji}\hat{\Psi}_{iL}\right)\left(\sum_i \tilde{\Omega}_{ji}\Psi_{iL}\right),$$

where $\tilde{\Omega}(j,:)$ represents the j-th row of matrix $\tilde{\Omega}$, and note that $\sum_i \tilde{\Omega}_{ji} = 1$. The covariance will be zero when sectoral exposure to labor, measured either by upstream markup (Ψ_{iL}) or wage pass-through ($\hat{\Psi}_{iL}$), is uniform across sectors. For instance, when firms freely reset their prices, wage pass-through is complete, $\hat{\Psi}_{iL} = 1$ for all i; alternatively, in absence of initial markups, the upstream markup is $\Psi_{iL} = 1$ for all i. Note also that when sector j corresponds to labor sector, $\tilde{\Omega}_{Li} = 0$ for all i by construction. The structure of production network also determines this covariance. For instance, in a vertical economy where sector j exclusively relies on a single input from sector i, this leads to $\tilde{\Omega}_{ji} = 1$ for sector i and $\tilde{\Omega}_{jk} = 0$ for all other sectors $k \neq i$, thereby resulting in a zero covariance. In general, the covariance $\mathrm{Cov}_{\tilde{\Omega}(j,:)}\left(\hat{\Psi}_{iL}, \Psi_{iL}\right)$ could be positive or negative, depending on initial markups, nominal rigidities and production networks.

3 Supply-Side Effect of Monetary Policy

Monetary policy changes the unit of account in the model economy and hence affects nominal wages and sectoral prices. When prices are fully flexible, monetary policy cannot affect resource reallocation and is thus neutral. However, when sectoral prices are sticky, wage pass-through to sectoral prices can vary across sectors in response to monetary policy shocks, leading to changes in relative sectoral prices. In an economy with initial distortions, relative price changes shift households' and downstream sectors' demand for goods produced by upstream sectors, and result in reallocation of labor and intermediate inputs across sectors, which in turn may lead to improvement in resource allocation. This section will investigate the conditions under which monetary policy improves total factor productivity (TFP) and the supply-side effect of monetary policy.

3.1 Total Factor Productivity

Our baseline model is a nested CES economy, as explored by Baqaee and Farhi (2020). The elasticity of substitution in production in sector i, $F_i(L_{i,\nu}, \{x_{ij,\nu}\}_{j=1}^{N})$, is σ_i. Since labor is the only primary input in our model, TFP is equivalent to labor productivity, $\mathrm{d}\log \mathrm{TFP} = \mathrm{d}\log Y - \mathrm{d}\log L$.

Theorem 1 provides a decomposition for TFP changes in response to changes in sectoral productivity $\mathrm{d}\log A_j$ and monetary policy. The expression (4) shows that a change in TFP can be broken down into two components: a direct technology channel and a misallocation channel. The direct technology channel is governed by the sum of distortion-adjusted ($\frac{\Psi_{jL}}{\Lambda_L}$) Domar-weighted ($\lambda_i$) sectoral productivity changes ($\mathrm{d}\log A_j$). The misallocation channel captures the impact of allocative efficiency through reallocation of resources across sectors in response to exogenous shocks, which critically depends on the aggregate wedge (captured by the average markup $1/\Lambda_L$), elasticity of substitution among inputs in both consumption and production (σ_j), the ratio of sectoral cost to GDP ($\lambda_j/\bar{\mu}_j$), and the covariance between sectoral price change ($\mathrm{d}\log p_i$) and upstream markup (Ψ_{iL}).

Theorem 1. *In response to sectoral productivity and monetary policy shocks, the change in TFP is governed by*

$$\mathrm{d}\log \mathrm{TFP} = \underbrace{\sum_{j=1}^{N} \lambda_j \frac{\Psi_{jL}}{\Lambda_L} \mathrm{d}\log A_j}_{\textit{Direct technology channel}} + \underbrace{\frac{1}{\Lambda_L} \sum_{j=0}^{N} \frac{\sigma_j \lambda_j}{\bar{\mu}_j} \mathrm{Cov}_{\tilde{\Omega}(j,:)} (\mathrm{d}\log p_i, \Psi_{iL})}_{\textit{Misallocation channel}} \tag{4}$$

with $\sigma_0 = \sigma_C$*,* $\lambda_0 = 1$*,* $\bar{\mu}_0 = 1$ *and* $\tilde{\Omega}(0,:) = b'$*.*

Theorem 1 demonstrates that sectoral productivity shocks influence TFP not only through the direct technology channel but also via the misallocation channel, as these shocks impact sectoral prices as well. In contrast, monetary policy affects TFP exclusively through the misallocation channel. Note that this misallocation channel is absent when all cross-sector elasticities of substitution are zero. The reason is that reallocation of labor and intermediate inputs across sectors is induced by relative changes in sectoral demands, which are caused by relative price changes in response to monetary policy shocks. If elasticities of substitution among sectoral products in consumption and production are zero ($\sigma_j = 0, \forall j$), relative price movements cannot change households and firms relative demand, and therefore, there is no reallocation of resources and consequently no change in TFP in response to monetary policy shocks.

If the government can completely offset all of monopolistic markups by using a set of industry-specific taxes/subsidies $\{\tau_i\}_{i=1}^{N}$, such that there are no markups in the initial equilibrium ($\bar{\mu} = \mathbf{1}$), then our model is degenerated to the frameworks of La'O and Tahbaz-Salehi (2022) and Rubbo (2023). In such a scenario, where the initial equilibrium is efficient ($\Psi_{iL} = \Lambda_L = 1, \forall i$), the misallocation channel disappears, and Theorem 1 reverts to Hulten's Theorem (Hulten, 1978). The reason is that although relative price changes in response to shocks alternate relative demand for goods produced by different sectors, marginal product of inputs in sectors facing higher demand is the same as that in sectors experiencing lower demand in equilibrium, and therefore these marginal products of inputs induced by the opposite shift of demand exactly cancel each other out based on the Envelop theorem. Consequently there is no additional gain from resource reallocation, even with nominal rigidities in place.

Nevertheless, when the initial allocation of resources is inefficient due to monopolistic markups, products from sectors with initial markups are under-supply compared to the socially optimal supply of goods. In an economy with production networks, this under-supply is further amplified via double marginalization through production chains. In the following analysis, we focus on how monetary policy affects resource allocation across sectors, and highlight under what conditions, monetary policy is able to improve total factor productivity.

3.2 Supply-Side Effect

As stated in section 2.4, the fact that the monetary authority controls the money supply is equivalent to its direct influence on nominal wages. So in this section, we consider only a change in monetary policy and treat a change in nominal wages $\mathrm{d}\log w$ as a monetary policy instrument instead, while assuming sectoral productivities unchanged. The supply-side effect of monetary policy is defined as the response of TFP to a change in monetary policy (see Definition 1).

Definition 1. *The supply-side effect of monetary policy is defined as the response of Total Factor Productivity to a change in monetary policy, given primary inputs and other exogenous shocks constant.*

Combining expressions for changes in TFP, labor income share, and wage pass-through to consumer price, the supply-side effect of monetary policy can be written as,

$$\frac{\mathrm{d}\log \mathrm{TFP}}{\mathrm{d}\log w} = 1 - \underbrace{\frac{\mathrm{d}\log P^Y}{\mathrm{d}\log w}}_{\hat{\Lambda}_L} - \underbrace{\frac{\mathrm{d}\log \Lambda_L}{\mathrm{d}\log w}}_{\doteq \xi} = 1 - \hat{\Lambda}_L - \xi,$$

where $\hat{\Lambda}_L$ stands for wage pass-through to consumer price, and ξ denotes elasticity of labor income share to nominal wages. Substituting wage path-through to sectoral price $\frac{\mathrm{d}\log p_i}{\mathrm{d}\log w} = \hat{\Psi}_{iL}$ into equation (4) in Theorem 1, we obtain one of key results in this study in Proposition 1. This proposition shows that the supply-side effect of monetary policy is driven by the misallocation channel, which can be further characterized by model primitives in terms of production networks, nominal rigidities, initial markups, and cross-sector elasticities of substitution.

Proposition 1. *The supply-side effect of monetary policy is given by the following sufficient statistic,*

$$1 - \hat{\Lambda}_L - \xi = \underbrace{\frac{1}{\Lambda_L}\sigma_C \mathrm{Cov}_b\left(\hat{\Psi}_{iL}, \Psi_{iL}\right)}_{\substack{\text{Reallocation due to}\\ \text{substitution in consumption}}} \quad + \quad \underbrace{\frac{1}{\Lambda_L}\sum_{j=1}^{N}\frac{\sigma_j \lambda_j}{\bar{\mu}_j}\mathrm{Cov}_{\tilde{\Omega}(j,:)}\left(\hat{\Psi}_{iL}, \Psi_{iL}\right)}_{\substack{\text{Reallocation due to}\\ \text{substitution in production}}}. \tag{5}$$

The supply-side effect in equation (5) depends on the average distortion, elasticities of substitution and the weighted covariance between wage pass-throughs ($\hat{\Psi}_{iL}$) and upstream markups (Ψ_{iL}) for consumer and each producer. When sectors with higher upstream markups tend to exhibit lower wage pass-throughs and hence stickier prices, and vice versa, sectors with higher upstream markups will raise their prices to a less extent than sectors with lower upstream markups in response to an expansionary monetary policy shock, implying that prices of higher upstream markup sectors become relatively cheaper than those of lower upstream markup sectors. A lower relative sectoral price leads to a higher demand and output. As a result, both labor and intermediate inputs will be re-allocated from sectors with lower upstream markups to sectors with higher upstream markups. Note that marginal product of input in sectors with higher upstream markups is greater than that in sectors with lower upstream markups. Consequently, this reallocation of resources leads to an improvement in TFP. Given positive elasticities of substitution and covariances, a more distorted economy (with a lower Λ_L) leads to a higher supply-side effect of monetary policy.

The use of sectoral output can be divided into two groups, consumption and downstream production. We label the supply-side effect from the former as the channel of reallocation due to substitution in consumption, while the latter as the channel of reallocation due to substitution in production. The reallocation channel from downstream production depends positively on elasticity of substitution, the ratio of sectoral cost to GDP, and a covariance term. Note that reallocation due to substitution in production only occurs in network economies. In an economy without input-output linkages, such reallocation mechanism is absent, and the supply-side effect of monetary policy is entirely attributed to substitution in consumption. In contrast to Baqaee et al. (2024), where a uniform initial markup across all firms nullifies the supply-side effect, our analysis reveals that the presence of input-output linkages introduces an additional layer of heterogeneity in the distribution of upstream distortions (Ψ_{iL}). Due to double marginalization along a supply chain and the fact that production processes are generally not symmetric across sectors, even when all firms in our economy have identical initial markups ($\bar{\mu}_i = \bar{\mu}^*, \forall i \in \mathcal{N}$), there still exists heterogeneity in upstream distortions, leading to an inefficient allocation of resources across sectors.

Another point worth of emphasizing is that the mechanisms of substitution in consumption and production operate quite differently due to different roles of labor in consumption and production. In consumption, the substitution is restricted among different sectoral products ($b_L = 0$), while in production, labor is a critical input ($\tilde{\Omega}_{iL} > 0, \forall i \in \mathcal{N}$), and the substitution occurs among intermediate inputs and between intermediate inputs and labor. To see this point clearly, we consider a scenario where sectoral prices are fully rigid (see Corollary 1). In such a scenario, there are no relative price changes among sectoral products, and accordingly, there is no reallocation due to substitution in consumption. Nevertheless, there is still a relative price change between intermediate inputs and labor ($\hat{\Psi}_{iL} = 0 < \hat{\Psi}_{LL} = 1, \forall i \in \mathcal{N}$), which allows for reallocation in production to generate a supply-side effect. Specifically, in response to an expansionary monetary shock, firms substitute labor for intermediate inputs since nominal wages rise while nominal sectoral prices remain unchanged. This reallocation towards intermediate goods enhances economy-wide allocative efficiency. This is because sectoral outputs are under-supply due to double marginalization ($\Psi_{iL} \leq \Psi_{LL} = 1, \forall i \in \mathcal{N}$), and taking use of more intermediate inputs leads to higher value added in the whole economy.

Corollary 1. *If sectoral prices are fully-rigid and initial markups are non-negative in all industries ($\theta_j = 0$ and $\bar{\mu}_j \geq 1$, $\forall j \in \mathcal{N}$), it follows that*

$$1 - \hat{\Lambda}_L - \xi = \underbrace{0}_{\substack{\text{Reallocation due to} \\ \text{substitution in consumption}}} + \underbrace{\frac{1}{\Lambda_L} \sum_{j=1}^{N} \sigma_j \lambda_j \tilde{\Omega}_{jL} (\bar{\mu}_j^{-1} - \Psi_{jL})}_{\substack{\text{Reallocation due to} \\ \text{substitution in production}}} \geq 0.$$

Equation (5) also shows that elasticity of substitution plays an essential role in determining the supply-side effect. Given the covariance terms unchanged, the more elastic of substitution among sectoral goods in consumption and/or among sectoral inputs in production, the larger the supply-side effect. The reason is that the demand from households and firms responds more strongly to relative sectoral price changes when the elasticity of substitution is higher. This implies that monetary policy has a greater impact on resource reallocation.

3.3 Illustrative Examples

To better understand how our model works, we explore three simple network economies to illustrate the decomposition in Proposition 1.

Example 1. A Vertical Economy

Consider first a vertical supply chain where the most upstream sector N produces good N using labor, while each of other sectors $i \neq N$ uses output produced by its immediate upstream sector $i+1$ as an intermediate input, with households consuming only output produced by sector 1. In this economy, there is only one feasible allocation of resources for a given level of labor supply, implying that the misallocation channel through monetary policy is absent.

One can verify that the sufficient statistic $1 - \hat{\Lambda}_L - \xi$ becomes zero since there is only one final good and one input involved in both consumption and sectoral production. Therefore the support of $\tilde{\Omega}(j,:)$ and b degenerates to a single point, which yields,

$$\mathrm{Cov}_b\left(\hat{\Psi}_{iL}, \Psi_{iL}\right) = 0 \quad \text{and} \quad \mathrm{Cov}_{\tilde{\Omega}(j,:)}\left(\hat{\Psi}_{iL}, \Psi_{iL}\right) = 0, \forall j \in \mathcal{N}.$$

Example 2. A Horizontal Economy

In a horizontal economy, each sector $i \in \mathcal{N}$ only uses labor to produce goods, which are sold directly to households. Given the absence of input-output linkages, the horizontal economy can be regarded as an economy with heterogeneous firms. When there is heterogeneity in markups, the cross-sector resource allocation is inefficient, providing an avenue for monetary policy to affect TFP. From Proposition 1, the sufficient statistic of supply-side effect of monetary policy can be simplified as,

$$1 - \hat{\Lambda}_L - \xi = \underbrace{\sigma_C \frac{\mathrm{Cov}_b(\theta, \bar{\mu}^{-1})}{\mathbb{E}_b(\bar{\mu}^{-1})}}_{\substack{\text{Reallocation due to} \\ \text{substitution in consumption}}} + \underbrace{0}_{\substack{\text{Reallocation due to} \\ \text{substitution in production}}}.$$

Note that there is no substitution in production since labor is the only input in each sector in this horizontal economy. The misallocation channel is driven entirely by substitution in consumption, characterized by the covariance between price rigidities (θ) and inverse of markups ($\bar{\mu}^{-1}$), using expenditure shares as weights. This result echoes the insight of Baqaee et al. (2024), which emphasizes that the response of TFP to monetary policy critically depends on the correlation between firms' markups and price rigidities in an economy with heterogeneous firms. In response to an expansionary monetary shock in our model, sectors with rigid prices (lower θ_i) increase their prices to a less extent than those with more flexible prices, leading to a reallocation of resources towards sectors with high price rigidities. If sectors with less flexible prices tend to have high initial markups, $\mathrm{Cov}_b(\theta, \bar{\mu}^{-1}) > 0$, monetary easing shifts resources from sectors with low markups to sectors with high markups, resulting in an improvement in allocative efficiency and therefore an increase in TFP. In addition, a higher average distortion (larger $1/\mathbb{E}_b(\bar{\mu}^{-1})$), given $Cov_b(\theta, \bar{\mu}^{-1})$ remains unchanged, further amplifies the supply-side effect.

Example 3. A Roundabout Economy

Now, consider a one-sector roundabout economy in which a representative firm combines labor and its own goods by using a CES production function with elasticity of substitution σ_1,

$$\frac{y_1}{\bar{y}_1} = A_1 \left[(1-\alpha) \left(\frac{x_{11}}{\bar{x}_{11}} \right)^{\frac{\sigma_1 - 1}{\sigma_1}} + \alpha \left(\frac{L_1}{\bar{L}_1} \right)^{\frac{\sigma_1 - 1}{\sigma_1}} \right]^{\frac{\sigma_1}{\sigma_1 - 1}},$$

where $\alpha \in (0, 1)$ represents the labor share in production. According to Proposition 1, it follows that:

$$1 - \hat{\Lambda}_L - \xi = \underbrace{0}_{\substack{\text{Reallocation due to} \\ \text{substitution in consumption}}} + \underbrace{(1-\alpha)\sigma_1 \frac{1-\theta_1}{1-(1-\alpha)\theta_1} \frac{1-\bar{\mu}_1^{-1}}{1-(1-\alpha)\bar{\mu}_1^{-1}}}_{\substack{\text{Reallocation due to} \\ \text{substitution in production}}}.$$

In contrast to the horizontal economy in Example 2, this roundabout economy has a single final good in the consumption basket, and therefore no reallocation occurs due to substitution in consumption. Nevertheless, a positive supply-side effect still arises from substitution in production when sectoral price is rigid ($\theta_1 < 1$), elasticity of substitution is positive ($\sigma_1 > 0$) and there exists an initial distortion ($\bar{\mu}_1 > 1$). In response to an expansionary monetary shock, firms substitute labor with intermediate inputs in production. This substitution occurs because wage pass-through is incomplete due to nominal rigidity ($\mathrm{d}\log p_1 < \mathrm{d}\log w$, or $\hat{\Psi}_{1L} = \frac{\alpha\theta_1}{1-(1-\alpha)\theta_1} < 1$). This reallocation of resources towards intermediate goods enhances allocative efficiency, since intermediate input is under deployed relative to the socially optimal level ($\Psi_{1L} = \frac{\alpha\bar{\mu}_1^{-1}}{1-(1-\alpha)\bar{\mu}_1^{-1}} < 1$).

Note also that the supply-side effect decreases in α. This is because wage pass-through becomes weaker (smaller $\hat{\Psi}_{1L}$) and upstream markup is larger (smaller Ψ_{1L}) when there exists more "roundaboutness" in production (lower α), resulting in a larger reallocation in production. Another point worth of emphasizing is that the supply-side effect is more pronounced when firms suffer more initial distortions (higher $\bar{\mu}_1$) or price rigidity (lower θ_1). The reason is that a stickier price is associated with a larger increase in real wages in response to an expansionary monetary policy, which in turn pushes firms more intensively taking use of intermediate inputs, and therefore, labor productivity rises. Following a similar logic, by using more intermediate inputs, an economy with a higher initial markup generates a higher gain from harvesting larger marginal product of input, and accordingly the supply-side effect becomes stronger.

4 Output Response and Phillips Curve

Monetary policy not only affects households' labor supply through the intratemporal decision between consumption and leisure, but may also change the allocation of labor and intermediate inputs across sectors. In this section, we will explore how monetary policy changes real output through its traditional New Keynesian demand-side effect and also the supply-side effect explored in the previous section. In addition, we will show that a positive supply-side effect can flatten the slopes of all sectoral Phillips curves.

4.1 Output Response

Given the specification of households' preference, the elasticity of real final output with respect to money supply is given by Proposition 2. It shows that the response of real output to a change in money supply can be decomposed into two components up to a first-order approximation: a supply-side effect that arises from a change in allocative efficiency, and a demand-side effect due to nominal price rigidities that has been studied in the New Keynesian models with production networks such as Pasten et al. (2020), La'O and Tahbaz-Salehi (2022), Afrouzi and Bhattarai (2023), Rubbo (2023) and Baqaee et al. (2024).

The demand-side effect arises from the endogenous response of labor supply to a change in money supply. When the monetary authority expands the money supply, nominal wages increase, but sectoral prices, along with consumer prices, might not respond to the same extent due to price rigidities. This incomplete wage pass-through to consumer price ($\hat{\Lambda}_L < 1$) results in a rise in real wages, which in turn leads households to supply more labor to firms, particularly when the wealth effect on labor supply is relatively weak (small γ). For the supply-side effect, Proposition 2 explicitly shows how TFP responds to monetary policy. When $1 - \hat{\Lambda}_L - \xi > 0$, an expansionary monetary policy increases total factor productivity, and vice versa.

Another point worth of emphasizing is that the Frisch elasticity of labor supply $1/\varphi$ also determines the magnitude of supply-side effect of monetary policy. We find that a less elastic labor supply (higher φ) enhances the supply-side effect. This is because with a less elastic labor supply, changes in nominal wages due to shocks are more pronounced, while the labor supply itself remains relatively unresponsive to these changes. In a model featuring price rigidities and initial markups, this larger response in nominal wages leads to more significant changes in ex-post markups, thus amplifying the supply-side effect. In an extreme case where labor supply becomes completely inelastic, as $\varphi \to \infty$, monetary policy has no impact on labor supply. In such a scenario, the supply-side effect is given by $\mathrm{d}\log \mathrm{TFP}/\mathrm{d}\log M = (1 - \hat{\Lambda}_L - \xi)/(1 - \xi)$, which is independent of households' preferences.

Proposition 2. *Following a monetary shock, the output response can be broken down into supply- and demand-side effects*

$$\frac{\mathrm{d}\log Y}{\mathrm{d}\log M} = \underbrace{\frac{\mathrm{d}\log \mathrm{TFP}}{\mathrm{d}\log M}}_{\textit{Supply-side effect}} + \underbrace{\frac{\mathrm{d}\log L}{\mathrm{d}\log M}}_{\textit{Demand-side effect}} = \frac{1 - \hat{\Lambda}_L + \varphi(1 - \hat{\Lambda}_L - \xi)}{1 + (\gamma - 1)\hat{\Lambda}_L + \varphi(1 - \xi)}, \tag{6}$$

specifically,

$$\frac{\mathrm{d}\log \mathrm{TFP}}{\mathrm{d}\log M} = \frac{(\gamma + \varphi)(1 - \hat{\Lambda}_L - \xi)}{1 + (\gamma - 1)\hat{\Lambda}_L + \varphi(1 - \xi)} \quad \textit{and} \quad \frac{\mathrm{d}\log L}{\mathrm{d}\log M} = \frac{1 - \hat{\Lambda}_L - \gamma(1 - \hat{\Lambda}_L - \xi)}{1 + (\gamma - 1)\hat{\Lambda}_L + \varphi(1 - \xi)}.$$

Our findings here are closely related to recent studies on production networks, but differ from the literature in several aspects. Specifically, compared to La'O and Tahbaz-Salehi (2022) which examines monetary non-neutrality in an efficient economy with input-output linkages, our demand-side effect replicates their Proposition 5 when we eliminate initial wedges. However, their model does not account for any supply-side effect, while our paper investigates both the supply-side and demand-side effects of monetary policy. Additionally, our work relates to Baqaee et al. (2024), which also decomposes the output response into supply- and demand-side effects. However, our approach differs from theirs in two aspects. First, the demand-side effect in our model is defined as the response of labor, which includes an adjustment resulting from the misallocation channel, $-\gamma(1-\hat{\Lambda}_L-\xi)$, whereas they attribute this component of labor response to the supply-side. Second, we examine the supply-side and demand-side effects of monetary policy in an economy with production networks, without relying on any real rigidities to generate significant supply-side effects.

4.2 The Divine Coincidence Condition

In an economy with initial distortions, the misallocation channel may have a first-order impact on TFP. Up to a first order approximation, the misallocation channel in equation (4) can be written as the difference between an output gap $\tilde{y}$ (the logarithmic difference between sticky-price and flexible-price equilibria) and an employment gap $\tilde{l}$, which can be further expressed as a weighted sum of sectoral price inflation π_k,

$$\tilde{y} - \tilde{l} = \sum_{k=1}^{N} \underbrace{\left(\frac{1}{\Lambda_L} \sum_{j=0}^{N} \frac{\sigma_j \lambda_j}{\bar{\mu}_j} \text{Cov}_{\tilde{\Omega}^{(j)}} \left(\tilde{\Psi}_{ik}, \Psi_{iL} \right) \right) (1 - \theta_k^{-1})}_{\doteq \mathcal{J}_k} \pi_k \doteq \mathcal{J}' \pi,$$

where $\mathcal{J}$ is an $N \times 1$ vector with the k-th component $\mathcal{J}_k$ measuring the impact of sector k's inflation on allocative efficiency. A larger $\mathcal{J}_k$ implies that a given level of sectoral inflation π_k could generate a greater improvement in allocative efficiency. $\mathcal{J}_k$ increases when substitution of inputs in production and substitution of varieties in consumption become more elastic, and the covariances between desired pass-throughs of sectoral inflation π_k to downstream sectors and upstream markups are greater in magnitude, in addition to stickier price and larger size in sector k.[9] Note also that when the initial allocation is efficient, the misallocation channel is absent ($\mathcal{J} = \mathbf{0}$).

From households labor supply and firms optimal conditions, output gap $\tilde{y}$ and sectoral price inflation rates π can be linked via a divine coincidence condition in Lemma 1, which reflects a trade-off between output gap and sectoral inflation without a cost-push term. The divine coincidence condition shows that when aggregate output exceeds that in the flexible price equilibrium, price inflation on average has to rise to make firms produce more output. Nevertheless, in a multi-sector economy, closing output gap $\tilde{y} = 0$ does not necessarily simultaneously stabilize price inflation in all sectors.

Lemma 1. *Assume no sector has fully rigid prices ($\theta_i \neq 0, \forall i$). The divine coincidence condition in*

[9] In a Cobb-Douglass economy, the big bracket in $\mathcal{J}_k$ can be simplified by using Lemma 8 in the Appendix as, $\frac{1}{\Lambda_L} \sum_{j=0}^{N} \frac{\lambda_j}{\bar{\mu}_j} \text{Cov}_{\tilde{\Omega}(j,:)} \left(\tilde{\Psi}_{ik}, \Psi_{iL} \right) = \lambda_k \frac{\Psi_{kL}}{\Lambda_L} - \tilde{\lambda}_k$. Then, $\mathcal{J}_k = (\lambda_k \frac{\Psi_{kL}}{\Lambda_L} - \tilde{\lambda}_k)(1 - \theta_k^{-1}) = \tilde{\lambda}_k (1 - \frac{\lambda_k}{\tilde{\lambda}_k} \frac{\Psi_{kL}}{\Lambda_L})(\theta_k^{-1} - 1)$.

a distorted economy is given by:

$$(\gamma + \varphi)\tilde{y} = [\tilde{\lambda}'(\Theta^{-1} - I) + \varphi\mathcal{J}']\pi. \tag{7}$$

There are two components in the coefficients of sectoral price inflation on the right hand side of equation (7). The first component $\tilde{\lambda}'(\Theta^{-1} - I)$ captures cost-based Domar-weighted sectoral price rigidities, as emphasized by Rubbo (2023), implying that for a given level of sectoral inflation, sectors with stickier prices and larger sizes have a greater impact on real output. The second component demonstrates a scenario where if sectoral price movements enhance allocative efficiency (i.e., $\mathcal{J}'\pi > 0$), sectoral inflation would further increase output. This effect is amplified by the inverse of Frisch elasticity, as the misallocation channel has more pronounced effects on output with a less elastic labor supply (see Proposition 2). In an extreme case where labor supply is fully elastic (as φ approaches zero), the misallocation channel becomes irrelevant to the divine coincidence condition and output response is independent of resource reallocation since any improvement in TFP due to reallocation is precisely offset by its adverse impact on labor supply.

4.3 Phillips Curves

Combining the divine coincidence condition with sectoral inflation $\pi = \hat{\Psi}^n_{(L)} \mathrm{d}\log w - \hat{\Psi}^n \Theta \mathrm{d}\log A$, we obtain a wage Phillips curve,

$$[1 - \hat{\Lambda}_L + \varphi(1 - \hat{\Lambda}_L - \xi)]\mathrm{d}\log w = (\gamma + \varphi)\tilde{y} + (\tilde{\lambda}' - \hat{\lambda}'\Theta + \varphi\mathcal{J}'\hat{\Psi}^n\Theta)\mathrm{d}\log A.$$

The coefficient of the change in nominal wages comprises two components: one reflects an increase in real wages $(1 - \hat{\Lambda}_L)$ due to nominal rigidities, and the other is driven by the misallocation channel. Rearranging the coefficient of wage inflation to the right-hand side of the equation above, we obtain the wage Phillips curve, which is flatter relative to the benchmark model without initial markups when resource allocation becomes more efficient following shocks and labor supply is not fully elastic. The reason is that in response to a positive output gap $\tilde{y}$, wages must rise to attract more labor to support additional output. However, enhanced allocative efficiency decreases firms' labor demand, thereby moderating the required increase in wages.

Note that sectoral value added is generated by labor through both direct labor input and indirect labor input from its upstream sectors. Therefore, a dampened wage inflation in response to output gap implies that firms adjust their product prices to a less extent when there exists a positive supply-side effect. Combining the wage Phillips curve with explicit expressions for sectoral inflation rates, Proposition 3 presents sectoral Phillips curves in the economy with initial distortions. It states that all sectoral price inflation Phillips curves become flattened when expansionary monetary policy improves allocative efficiency.

Proposition 3. *Sectoral Phillips curves in a distorted economy are given by*

$$\underbrace{\pi}_{N\times 1} = \underbrace{\mathcal{K}}_{N\times 1}\tilde{y} + \underbrace{\mathcal{V}}_{N\times N}\underbrace{\mathrm{d}\log A}_{N\times 1} \tag{8}$$

where $\mathcal{K}$ and $\mathcal{V}$ denote the slope and residual coefficient matrix of the Phillips curves, respectively. Specifically, $\mathcal{K}$ is an $N \times 1$ vector given by

$$\mathcal{K} = \frac{\gamma + \varphi}{1 - \hat{\Lambda}_L + \varphi(1 - \hat{\Lambda}_L - \xi)}\hat{\Psi}^n_{(L)},$$

and $\mathcal{V}$ is an $N \times N$ matrix defined as

$$\mathcal{V} = \frac{1}{1 - \hat{\Lambda}_L + \varphi(1 - \hat{\Lambda}_L - \xi)}\hat{\Psi}^n_{(L)}(\tilde{\lambda}' - \hat{\lambda}'\Theta + \varphi\mathcal{J}'\hat{\Psi}^n\Theta) - \hat{\Psi}^n\Theta.$$

When the initial equilibrium is efficient, sectoral Phillips curves are then in line with Proposition 2 in Rubbo (2023). Input-output linkages make sectoral Phillips curves flatter through compounded price rigidities. In our model economy with initial distortions, sectoral Phillips curves are further flattened due to improvement in allocative efficiency. The reason is that when monetary expansion improves allocative efficiency, the supply-side effect brings up higher output but without much increase in inflation. Following a similar logic, other aggregate Phillips curves, including consumer price Phillips curve, which uses expenditure shares as weights, also flatten due to the supply-side effect of monetary policy. We summarize these results in Corollary 2.

Corollary 2. *In a distorted economy, wage, sectoral, and aggregate Phillips curves become flatter if an expansionary monetary policy improves allocative efficiency.*

5 Optimal Monetary Policy

The divine coincidence condition in Lemma 1 reveals that sectoral price inflation and output gap may not be stabilized simultaneously. Additionally, the presence of a supply-side effect of monetary policy allows the monetary authority to stimulate the economy to improve allocative efficiency, which in turn results in an inflation bias. This section will explore the optimal monetary policy.

5.1 Welfare Function

Our model economy features two frictions: nominal rigidities and initial markups. On the one hand, higher price dispersions both within and across sectors due to nominal rigidities reduce welfare (see, for instance, La'O and Tahbaz-Salehi, 2022; Rubbo, 2023). On the other hand, an economy with initial markups produces lower output than that without markups, and therefore, the monetary authority may use its policy instruments to boost up output as in Galí (2015). Proposition 4 below provides a second-order approximation of households welfare around a distorted flexible price equilibrium. The welfare function in equation (9) comprises five terms: a first-order bias, output gap volatility, within-sector and cross-sector price dispersions, and variations in allocative efficiency.

Proposition 4. *Under the assumption of small distortions, up to second-order approximation, the welfare function is given by*

$$\mathbb{W} = \underbrace{(1-\Lambda_L)\tilde{y} + \Lambda_L \mathcal{J}'\pi}_{\textit{First-order bias}} \underbrace{-\frac{\gamma+\varphi}{2}(\tilde{y} - \frac{1+\varphi}{\gamma+\varphi}\mathcal{J}'\pi)^2}_{\textit{Volatility of output gap}} \underbrace{-\frac{1}{2}\pi'\mathcal{H}_1\pi}_{\textit{Within-sector price dispersion}}$$
$$\underbrace{-\frac{1}{2}\pi'\mathcal{H}_2\pi}_{\textit{Cross-sector price dispersion}} \underbrace{-\frac{\gamma-1}{2}\frac{1+\varphi}{\gamma+\varphi}\pi'\mathcal{J}\mathcal{J}'\pi}_{\textit{Variation in allocative efficiency}} \tag{9}$$

where $\mathcal{H}_1 = \mathrm{diag}(\epsilon)\mathrm{diag}\left((\theta^{-1}-1)\circ\lambda\circ\frac{\Psi_{(L)}}{\Lambda_L} - \mathcal{J}\right)$ *and* $\mathcal{H}_2 = (I-\Theta^{-1})\mathcal{B}(I-\Theta^{-1})$ *are* $N\times N$ *matrices, with* $\mathcal{B}$ *defined as an* $N\times N$ *matrix where each element* $\mathcal{B}(k,l)$ *equals* $\sum_j \sigma_j\lambda_j \mathrm{Cov}_{\Omega(j,:)}(\Psi_{ik},\Psi_{il})$.

The first term in equation (9) represents a first-order bias, characterized by the aggregate wedge weighted sum of output gap $\tilde{y}$ and allocative efficiency $\mathcal{J}'\pi$. The weight on output gap, $1-\Lambda_L$, measures the aggregate wedge between natural and efficient levels of output. When some sectors exhibit positive markups ($\bar{\mu}_j > 1$ for some j, and $\bar{\mu}_i \geq 1$ for others), these markups propagate through production networks and lead to a positive aggregate wedge. Monetary policy could obtain first-order welfare gains by driving the economy up towards its efficient production frontier.

The second term illustrates a welfare loss due to output volatility, captured by an effective output gap, $\tilde{y} - \frac{1+\varphi}{\gamma+\varphi}\mathcal{J}'\pi$. This effective output gap is a weighted average of output and employment gaps, reflecting a trade-off between closing these two gaps when they diverge. When an expansionary monetary policy is able to improve allocative efficiency, $\mathcal{J}'\pi > 0$, output in the economy will increase but without much welfare loss caused by inflation. Therefore, output should be stabilized around an efficiency-adjusted target.

The third and fourth terms in the welfare function characterize welfare losses due to within-sector and cross-sector price dispersions, respectively. Within each sector, firms with and without the opportunity to adjust prices in response to shocks may experience different prices. This price dispersion leads to within-sector distortions since resources are inefficiently allocated to firms with lower prices. Similarly, sectoral prices may not adjust uniformly to changes in input costs due to nominal rigidities and input-output linkages, resulting in lower welfare. This cross-sector price dispersion contributes to a second-order misallocation effect (La'O and Tahbaz-Salehi, 2022 and Rubbo, 2023). Specifically, in a distorted horizontal economy, the cross-sector price dispersion can be written as,

$$\frac{1}{2}\pi'\mathcal{H}_2\pi = \frac{1}{2}\sigma_C \text{Var}_\beta(\text{d} \log \mu_i),$$

where β denotes a vector of weights in the variance operator, with element $\beta_k = b_k\bar{\mu}_k^{-1}/\mathbb{E}_b(\bar{\mu}_k^{-1})$. The right hand side of the expression above is a measure of second-order misallocation effect as in Hsieh and Klenow (2009).[10] Higher price dispersion across sectors results in lower TFP and welfare.

The last term in equation (9) shows welfare cost due to variations in allocative efficiency. A key parameter that determines the sign of this term is γ, which captures the wealth effect on labor supply. When resource allocation improves in response to shocks, $\mathcal{J}'\pi > 0$, TFP increases accordingly, resulting a higher level of output and households income, given other conditions unchanged. If the wealth effect on labor is strong, $\gamma > 1$, higher income from a better resource allocation dampens the response of labor supply, resulting in lower welfare, and vice versa. The knife-edge case is when $\gamma = 1$, where this variation in allocative efficiency does not generate any welfare loss.

5.2 Optimal Policy

Optimal monetary policy maximizes the welfare function $\mathbb{W}$ in Proposition 4 by choosing output gap and sectoral inflation rates, subject to sectoral Phillips curves in Proposition 3.

[10] In Hsieh and Klenow (2009), the negative effect of distortions on TFP is summarized by the variance of log TFPR: $\frac{1}{2}\sigma_C\text{Var}(\text{d} \log \text{TFPR}_i)$. Under constant returns to scale, changes in TFPR coincide with changes in markups: $\text{d} \log \text{TFPR}_i = \text{d} \log \mu_i$.

In a multi-sector model with input-output linkages and various frictions, a central bank equipped with a single policy instrument generally isn't able to close all of gaps shown in the welfare function. On one hand, the central bank must balance the output gap, sectoral inflation rates, and allocative efficiency, all of which contribute to second-order welfare losses. On the other hand, in an economy with initial distortions, the central bank encounters an additional trade-off: balancing second-order welfare losses against first-order welfare gains. This additional trade-off results in an inflation bias. Proposition 5 presents the condition for sectoral inflation rates under the optimal monetary policy.

Proposition 5. *The optimal monetary policy is an inflation stabilization policy with a bias, which is determined by the following condition,*

$$\left[\underbrace{\tilde{\lambda}'(\Theta^{-1} - I) - \mathcal{J}'}_{\textit{Output gap}} + \underbrace{\frac{\gamma + \varphi}{\xi}(\hat{\Psi}^n_{(L)})'\mathcal{H}}_{\textit{Price dispersion}} + \underbrace{\frac{(1+\varphi)(\gamma - 1)(1 - \hat{\Lambda}_L - \xi)}{\xi}\mathcal{J}'}_{\textit{Variation in allocative efficiency}}\right]\pi$$
$$= \underbrace{1 - \Lambda_L}_{\textit{Inflation bias due to aggregate wedge}} + \underbrace{[1 + \varphi + (\gamma - 1)\Lambda_L](1 - \hat{\Lambda}_L - \xi)/\xi}_{\textit{Inflation bias due to supply-side effect}} \qquad (10)$$

where $\mathcal{H} = \mathcal{H}_1 + \mathcal{H}_2$.

The right-hand side of equation (10) highlights the source of the inflation bias, which arises from an aggregate wedge $1-\Lambda_L$ and the supply-side effect of monetary policy $1-\hat{\Lambda}_L-\xi$. In a distorted economy, the natural level of output is lower than its efficient level ($1-\Lambda_L > 0$), as in standard New Keynesian models. This discrepancy due to the aggregate wedge provides room for the monetary authority to boost up output toward its efficient level via higher inflation. The other source of inflation bias comes from the supply-side effect, which is novel in our model. When monetary policy is able to generate a positive supply-side effect ($1 - \hat{\Lambda}_L - \xi > 0$), the monetary authority utilizes this effect to enhance allocative efficiency and welfare.

In contrast, the left-hand side of equation (10) illustrates how optimal policy balances various sources of second-order welfare losses and stabilizes an index of sectoral inflation rates. The weight on sectoral inflation in the index consists of three components, which minimize output gap, price dispersions within and across sectors, and variations in allocative efficiency, respectively. These weights are shaped by production networks, nominal rigidities, initial markups, elasticities of substitution, and preference parameters. A higher weight implies a lower desired inflation under the optimal policy.

As emphasized by La'O and Tahbaz-Salehi (2022) and Rubbo (2023), the weights of stabilizing output gap generally carry similar sectoral weights in the inflation index under the optimal monetary policy. These weights in equation (10) show that the monetary authority should assign higher weights to the following sectors in the inflation index: (i) sectors with stickier prices (lower θ_i) due to their higher price distortions; (ii) larger sectors, as indicated by higher cost-based Domar weights ($\tilde{\lambda}_i$), which have a greater impact on the economy for a given price dispersion; and (iii) sectors with less distortions (lower $\mathcal{J}_k$), whose price fluctuations in response to shocks have a limited impact on resource allocation.[11] Comparing with La'O and Tahbaz-Salehi (2022) and Rubbo (2023), point (iii) is novel in our study and it shows that the monetary authority assigns a lower weight to sectoral inflation that generates higher allocative efficiency to the whole economy (e.g. higher $\mathcal{J}_k$).

5.3 Alternative Stabilization Policies

In practice, the monetary authority may follow alternative stabilization policies to implement its monetary policy. In line with the literature, we consider two simple alternative rules: a consumer price inflation targeting rule, $\sum_{i=1}^{N} b_i \pi_i = 0$ and an output gap targeting rule, $\tilde{y} - \frac{1+\varphi}{\gamma+\varphi}\mathcal{J}'\pi = 0$, which is taken from the welfare function (9). Combining with the divine coincidence condition (7), this output-gap stabilization policy can be rewritten as an inflation-stabilization policy, $\sum_i \phi_i^{o.g.} \pi_i = 0$, with $\phi_i^{o.g.} = (1/\theta_i - 1)\tilde{\lambda}_i - \mathcal{J}_i$. Note that $b_i \neq \phi_i^{o.g.}$ generally, as the output-gap stabilization policy partly takes into account input-output linkages and adjustments in nominal rigidities and allocative efficiency. There are two main differences between the optimal monetary policy and these two simple rules. On one hand, optimal monetary policy creates an inflation bias due to the aggregate wedge and supply-side effects, which these simple rules do not address. On the other hand, these rules fail to consider price dispersion both within and across sectors, as well as variations in allocative efficiency.

6 Quantitative Analysis

In the previous sections, we show that monetary policy can influence realized sectoral markups in general, leading to resource reallocation across sectors and improvements in social welfare. Policymakers could take advantage of this supply-side effect to boost up TFP

[11] Appendix F presents an explicit expression for each component in a Cobb-Douglass economy.

by improving allocative efficiency of the economy, beyond merely stabilizing output and inflation as in standard New Keynesian models. A critical question is how important the supply-side effect will be when we take our model to data. In this section, we calibrate our model using data from the United States, quantitatively showing that both the supply-side and demand-side effects of monetary policy are essential in the transmission of monetary policy. Moreover, we find that these effects decline substantially in an economy without input-output linkages.

6.1 Calibration

Our model period is one quarter and includes two sets of parameters. The first set includes preference and production parameters, which are set to be standard values in the literature. Specifically, the relative risk aversion coefficient is $\gamma = 1$, and the inverse Frisch elasticity is $\varphi = 2$. In our baseline calibration, we set all within-sector elasticities of substitution at 6 (i.e., $\varepsilon_i = 6, \forall i \in \mathcal{N}$).[12] We also specify the elasticity of substitution in consumption to be one ($\sigma_C = 1$) and cross-sector elasticities of substitution in firms' production to be 0.5 (i.e., $\sigma_i = 0.5, \forall i \in \mathcal{N}$) following the literature Atalay (2017), Levchenko et al. (2019) and Devereux et al. (2023).

The second set of parameters includes input-output linkages, nominal rigidities, initial markups, final consumption shares, and productivity shocks. To estimate these parameters, we utilize four different but consistent datasets. First, we calibrate the cost-based input-output matrix $\tilde{\Omega}$ and consumption shares b using annual input-output data from the Bureau of Economic Analysis (BEA) in the United States. Following standard practice in the literature (see Baqaee and Farhi (2020) and La'O and Tahbaz-Salehi (2022)), we exclude sectors related to federal, state, and local government because our model does not consider the role of government. The dataset encompasses 66 industries. Since sectoral production technology hardly changes within a year, we assume that the quarterly cost-based input-output matrix $\tilde{\Omega}$ and consumption shares b keep constant throughout the year, matching their annual values.

We incorporate annual sectoral markups estimated by Baqaee and Farhi (2020) as one of the key parameters in our model. The data on markups, spanning from 1997 to 2015, include information on 66 industries in our input-output dataset. We use the user-cost (UC) markup series as our benchmark calibration for initial markups. For simplicity, we assume

[12] We do not directly calibrate within-sector elasticities to match initial markups in each sector and each period. Instead, we align the values of within-sector elasticities with benchmarks in the New Keynesian literature (e.g., McKay et al., 2016), and vary sectoral tax (subsidy) rates to match the gap between markups in the data and monopolistic markups in the model.

that quarterly markups remain constant within a year. By combining the estimated markups with the cost-based input-output matrices from the BEA, we construct the revenue-based input-output matrix Ω in our model.

Sectoral price rigidities are calibrated using data from Pasten et al. (2020) who utilize confidential micro-data from the Bureau of Labor Statistics (BLS) Producer Price Index (PPI) to analyze the frequency of price adjustments across different industries. The data covers the period from 2005 to 2011. This measure calculates the ratio of the number of price changes to the number of months in the sample. By merging the BEA input-output data with price adjustment data, classified according to the 3-digit codes of the North American Industry Classification System (NAICS), we calibrate price rigidity for each industry. Combining constant sectoral price rigidities with a time-varying cost-based input-output matrix allows us to derive a time-varying rigidity-adjusted input-output matrix $\hat{\Omega}$.

Finally, we measure productivity shocks by the sector-level growth rate of the Multifactor Productivity (MFP) index. The data for this analysis, covering the period from 1987 to 2019, is taken from the BEA/BLS Integrated Industry-Level Production Account (ILPA). The BEA input-output data are more disaggregated compared to the ILPA data. For each of the 66 industries, we calculate productivity shocks by disaggregating larger industries into smaller sub-industries. Annual productivities are linearly interpolated to generate quarterly values for each industry, which are then detrended to construct the covariance matrix of detrended productivity shocks.

6.2 Supply-Side Effect and Inflation Bias

Figure 1 provides a breakdown of the supply-side effect of monetary policy on TFP, the sufficient statistic $\mathrm{d}\log \mathrm{TFP}/\mathrm{d}\log w$ in equation (5). The solid blue line represents the supply-side effect due to substitution in consumption, the dashed purple line denotes the supply-side effect resulting from substitution in production, and the dotted red line with circle markers shows the sum of both blue and purple lines. Panel (a) reports numerical results based on our benchmark calibration. The supply-side effect (dotted red line) fluctuates between 0.005% and 0.031%, with an average of 0.018%, implying that a 1% increase in nominal wages stemming from expansionary monetary policy boosts up TFP by about 0.018%. The Panel shows that the reallocation from substitution in production co-moves quite closely with the total supply-side effect, accounting for 79% to 95% of the total effect in the data sample, while the reallocation due to substitution in consumption is quite small and remains stable over time.[13]

To better see how each sector contributes to the supply-side effect, Figure A.1 in the Appendix presents the contributions from households and 66 different sectors in 2006. The supply-side effect due to reallocation from consumption (households) accounts for 5.7% of the total effect, while some key sectors including construction, wholesale trade, and professional, scientific, and technical services contribute to about one-fifth of the total supply-side effect due to reallocation in production.

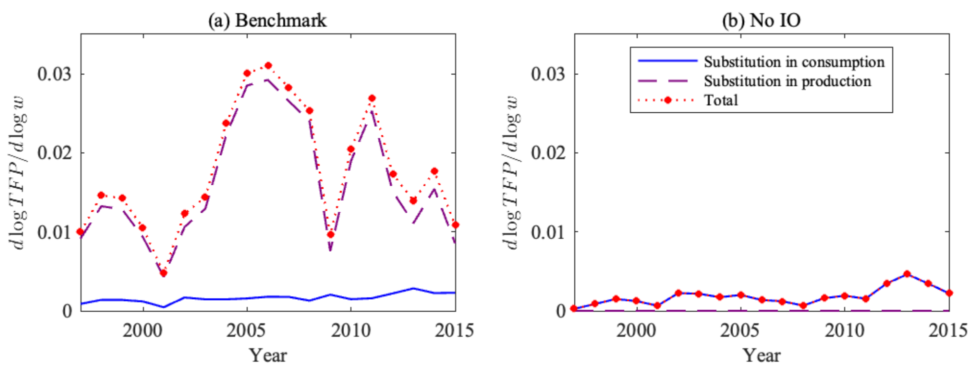

Figure 1: The Decomposition of Supply-Side Effect with UC Markups. Notes: This figure reports numerical values for two components in the supply-side effect of monetary policy. The solid blue line corresponds to the consumption-related reallocation channel, the dashed purple line shows the production-related reallocation channel, and the dotted red line with circle markers indicates the total effect.

Note that sectoral markups in the data substantially vary with business cycles. Firms typically experience lower markups during recessions and higher markups in economic booms. Panel (a) shows that the channel of reallocation from substitution in production varies significantly over time, with several troughs occurring in 2001, 2009, and 2015. During these years, some industries experienced severe markdowns, which were attributed to dramatic contraction in aggregate demand in 2001 and 2009, and substantial declines in energy commodity prices in 2015. This reduction in markups leads to a decrease in the total supply-side effect during economic downturns.

The analysis above shows that the supply-side effect of monetary policy is quantitatively important when we take our model to data. A next question is whether the monetary authority has an incentive to create inflation to harvest this supply-side effect and moreover whether the inflation bias induced by the supply-side effect is quantitatively important.

[13] Figure A.2 in the appendix shows that the correlation between wage pass-throughs to prices and upstream markups are usually positive in the data.

Figure 2 shows the inflation bias under the optimal monetary policy in equation (10). The dashed red line represents the inflation bias caused by the aggregate wedge, while the solid blue line shows the inflation bias arising from the supply-side effect. Panel (a) is for the baseline model. It shows that the inflation bias attributed to the supply-side effect moves quite along with that due to the aggregate wedge, fluctuating from 0.08% during economic downturns to 0.54% in booms, with an average around 0.30%.

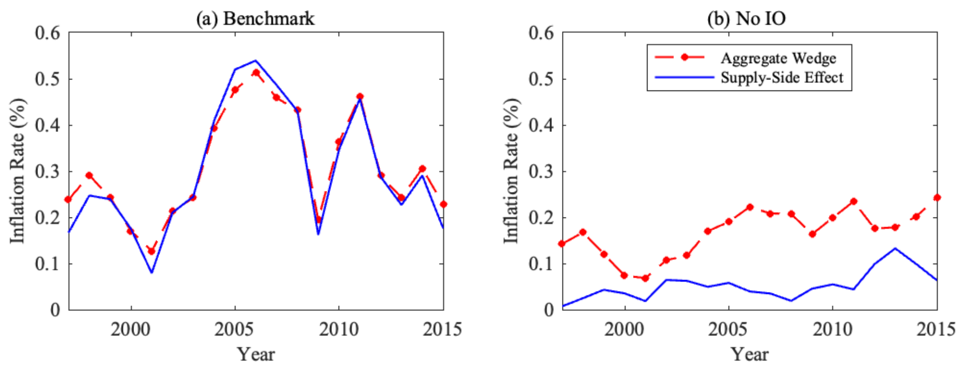

Figure 2: The Decomposition of Inflation Bias with UC Markups. Notes: This figure reports numerical values for two components on the right-hand side of equation (10) when the sum of industry weights on the left-hand side are normalized to 1. The blue line stands for the supply-side effect and the dashed red line with circle markers represents for the aggregate wedge.

What are the welfare implications under the optimal policy? Table 2 presents the overall welfare gains and its decomposition in equation (9). In our model, monetary policy affects output in two distinct ways. On one hand, expansionary monetary policy could push up output closer to its efficient level via creating higher inflation when there exists an aggregate wedge in the economy. On the other hand, it may also improve allocative efficiency and expand final output through the reallocation of labor and intermediate inputs when there exists a supply-side effect of monetary policy. However, it's important to note that expansionary monetary policy can also lead to higher inflation, potentially increasing price dispersion within and across sectors, and thus resulting in larger welfare losses.

Column (1) in Table 2 reports welfare losses (indicated by negative numbers) under the optimal monetary policy. The overall welfare loss is equivalent to 0.562% reduction in consumption relative to that under the flexible price equilibrium. However, both first-order welfare gains due to the aggregate wedge and the supply-side effect are positive and have similar magnitudes. The middle panel of Table 2 displays second-order welfare losses resulting from output gap volatility, as well as price dispersions within and across sectors. The within-sector price dispersion generates a welfare loss of 0.567% of consumption and the cross-sector price dispersion also give rise to a 0.094% reduction in consumption. In our baseline specification, we set $\gamma = 1$ and therefore the welfare cost associated with variation in allocative efficiency becomes zero.[14]

Table 2: Welfare Gains under Various Policies with UC Markups

	(1) Optimal Policy	(2) Output-Gap Targeting	(3) CPI Targeting
Welfare Gain (% of Real GDP)	-0.562	-0.617	-0.622
First order welfare gain			
Aggregate wedge	0.059	10^{-7}	0.001
Supply-side effect	0.048	10^{-6}	10^{-5}
Second order welfare gain			
Output gap volatility	-0.007	0	-0.002
Within-sector price dispersion	-0.567	-0.529	-0.529
Cross-sector price dispersion	-0.094	-0.088	-0.092
Variation in allocative efficiency	0	0	0
Cosine similarity to optimal policy	1	0.987	0.663

6.3 Production Networks and Supply-Side Effect

Our theory suggests that the share of intermediate inputs in production plays a critical role in determining the supply-side effect of monetary policy. How do changes in input-output linkages quantitatively affect the two reallocation channels of the supply-side effect, as well as the demand-side effect of monetary policy? Panel (b) of Figure 1 presents the same decomposition as in Panel (a) but firms do not employ any intermediate input in production. The model economy now is essentially degenerated to a horizontal economy. As shown in Section 3.3, the supply-side effect is completely driven by the reallocation from substitution in consumption as in Baqaee et al. (2024). The figure shows that the supply-side effect is positive, but it becomes much smaller than that in Panel (a). This comparison indicates that an economy with production networks as in the data is able to generate a sizable supply-side effect of monetary policy even without resorting to real rigidities as in Baqaee et al. (2024). Panel (b) in Figure 2 displays the inflation bias due to the aggregate wedge and the supply-side effect of monetary policy without input-output linkages in production. Both the aggregate wedge and the supply-side effect lead to a smaller inflation bias in the horizontal economy than the baseline model. More importantly, the inflation bias due to aggregate wedge significantly dominates that brought by the supply-side effect under

[14]Even when γ significantly deviates from 1, for example $\gamma = 2$, the welfare loss from variations in allocative efficiency remains negligible (smaller than 10^{-4}).

the optimal monetary policy, accounting for 57% to 95% of the total inflation bias in the data sample. Furthermore, Figure A.3 in the appendix demonstrates that the demand-side effect of monetary policy also significantly declines in the absence of input-output linkages, aligning with findings by Nakamura and Steinsson (2010) and Pasten et al. (2020).

6.4 Simple Rules of Monetary Policy

Can the monetary authority take use of alternative simple rules to approximate the optimal monetary policy? In this section, we examine two alternative simple rules of monetary policy. One is an output gap stabilization rule $\sum_i \phi_i^{o.g.} \pi_i = 0$ similar to Rubbo (2023), and the other is a CPI inflation targeting rule $\sum_{i=1}^N b_i \pi_i = 0$. Note that these two simple rules do not have the first-order inflation bias as the optimal monetary policy (equation (10)). Consequently, unlike the optimal monetary policy which trades off first-order welfare gains against second-order welfare losses, these two simple rules mainly balance second-order welfare losses induced by output gap, within and cross-sector price distortions, and variations in allocative efficiency.

Column (2) in Table 2 presents welfare gains and their decomposition under the output gap stabilization rule. Results show that the first order welfare gains are essentially zero, but welfare losses from within-sector and cross-sector price dispersions are smaller than those under the optimal monetary policy, although the cosine similarity (see La'O and Tahbaz-Salehi (2022)) of weights on sectoral inflation rates between optimal monetary policy and output stabilization is quite large, 0.987, implying that the distribution of weights on sectoral inflation is quite similar under these two monetary policies.

Column (3) in Table 2 presents welfare gains and their decomposition under the CPI inflation targeting rule. Welfare gains from first-order aggregate wedge and the supply-side effect are quite small, while welfare losses due to price dispersion across sectors are larger than those under output stabilization, and output gap variation also brings welfare losses. Comparing with output stabilization, CPI inflation targeting results in an even larger welfare loss. The reason is that output stabilization already partially accounts for output change due to allocative efficiency, and stabilizing such an output gap reduces welfare losses. The bottom line in Table 2 shows that sectoral weights under CPI inflation targeting are less aligned with those under the optimal policy compared to those under output gap stabilization.

6.5 Robustness Analysis

Figure 3 illustrates the inflation bias arising from the aggregate wedge and the supply-side effect of monetary policy under different preference parameterization. Panel (a) shows that an increase in wealth effect ($\gamma = 2$) amplifies the supply-side effect, leading to a higher inflation bias compared to the aggregate wedge, while a decrease in wealth effect ($\gamma = 0.5$) leads to a lower inflation bias induced by the supply-side effect. In addition, Proposition 5 suggests that the inflation bias due to the supply-side effect is also amplified with lower labor supply elasticity. Panels (c) and (d) confirm this point: a more elastic labor supply ($\varphi = 0.5$) leads to a lower inflation bias generated by the supply-side effect, while a less elastic supply ($\varphi = 5$) enhances the impact of supply-side effect on inflation bias.

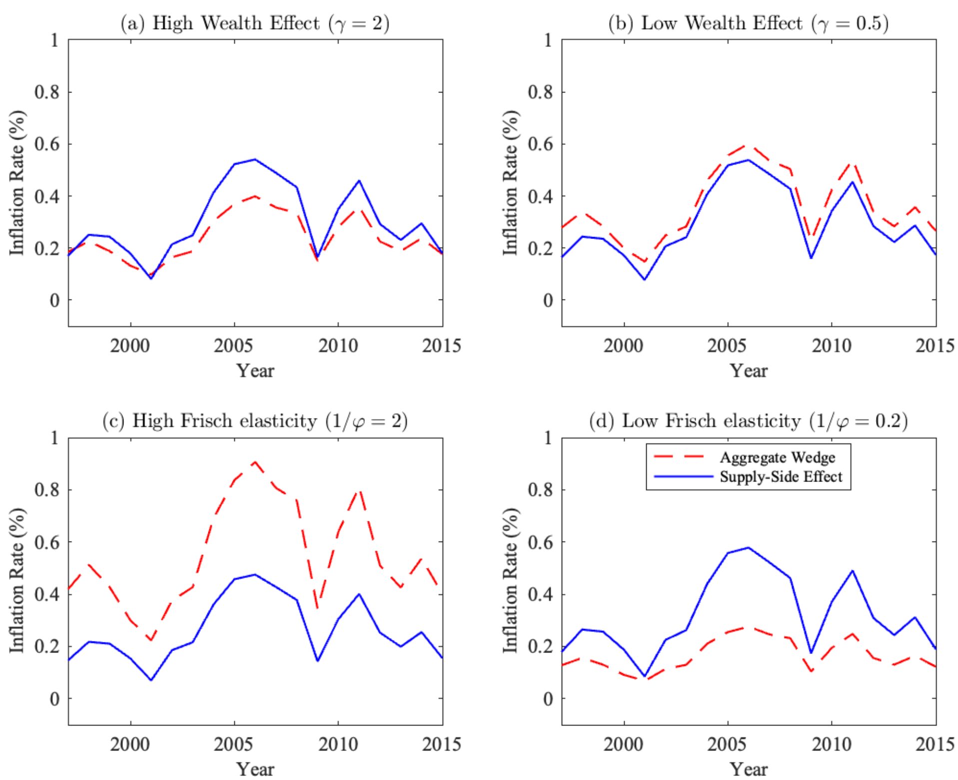

Figure 3: Decomposition of Inflation Bias with UC Markups, Different Preference Parameters

Heterogeneity in sectoral markups and input-output linkages might affect the supply-side effect as well. We then conduct two counterfactual experiments. First, we eliminate heterogeneity in sectoral markups by setting all sectors to have the same average markup within a year, while still allowing these markups to vary over time. The results in panel (a) of Figure A.4 are similar to the baseline calibration. In the second experiment, the input-output linkages for all years are adjusted to align with those from the first year of the data sample, while sectoral markups remain consistent with the data. Note that input-output linkages largely depend on the technology of production, which evolves slowly over time in the data. Results in panel (b) of Figure A.4 indicate that the supply-side effect is very similar to the baseline model.

7 Conclusion

This paper studies the supply-side and demand-side effects of monetary policy in a multi-sector economy with input-output linkages. We show that the supply-side effect of monetary policy is driven by two reallocation channels, reallocation due to substitution in consumption and reallocation due to substitution in production. These two channels become more pronounced when the elasticities of substitution are higher, sectors with larger upstream markups tend to exhibit lower wage pass-throughs, and aggregate wedge is larger.

The wage, sectoral, and aggregate Phillips curves become flatter when expansionary monetary policy improves allocative efficiency. Under the optimal monetary policy, the monetary authority has an incentive to inflate the economy due to both an aggregate wedge and the supply-side effect, and it assigns greater weights of inflation to larger, stickier and less distorted sectors.

When calibrating our model to data from the United States, we find that the supply-side effect of monetary policy is quantitatively important, and production networks play a crucial role in determining this effect. Under optimal monetary policy, the welfare gains from supply-side effects are comparable to those resulting from the aggregate wedge. Our sensitivity analysis indicates that our findings remain robust to variation in model parameterization. When we eliminate production networks, the supply-side and demand-side effects of monetary policy decline substantially.

References

Acemoglu, Daron, and Alireza Tahbaz-Salehi (2020) 'Firms, failures, and fluctuations: the macroeconomics of supply chain disruptions.' Technical Report, National Bureau of Economic Research

— (2024) 'The macroeconomics of supply chain disruptions.' *Review of Economic Studies* p. rdae038

Acemoglu, Daron, and Pablo D. Azar (2020) 'Endogenous production networks.' *Econometrica* 88(1), 33–82

Acemoglu, Daron, Asuman Ozdaglar, and Alireza Tahbaz-Salehi (2017) 'Microeconomic origins of macroeconomic tail risks.' *American Economic Review* 107(1), 54–108

Acemoglu, Daron, Ufuk Akcigit, and William Kerr (2015) 'Networks and the macroeconomy: An empirical exploration.' In 'NBER Macroeconomics Annual 2015, Volume 30' NBER Chapters (National Bureau of Economic Research, Inc) pp. 276–335

Acemoglu, Daron, Vasco M. Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi (2012) 'The network origins of aggregate fluctuations.' *Econometrica* 80(5), 1977–2016

Adão, Bernardino, Isabel Correia, and Pedro Teles (2003) 'Gaps and Triangles.' *The Review of Economic Studies* 70(4), 699–713

Afrouzi, Hassan, and Saroj Bhattarai (2023) 'Inflation and GDP dynamics in production networks: A sufficient statistics approach.' NBER Working Papers 31218, National Bureau of Economic Research, Inc, May

Altinoglu, Levent (2021) 'The origins of aggregate fluctuations in a credit network economy.' *Journal of Monetary Economics* 117(C), 316–334

Angeletos, George-Marios, and Jennifer La'O (2020) 'Optimal monetary policy with informational frictions.' *Journal of Political Economy* 128(3), 1027–1064

Atalay, Enghin (2017) 'How important are sectoral shocks?' *American Economic Journal: Macroeconomics* 9(4), 254–280

Baqaee, David, and Elisa Rubbo (2023) 'Micro propagation and macro aggregation.' *Annual Review of Economics* 15(Volume 15, 2023), 91–123

Baqaee, David R. (2018) 'Cascading failures in production networks.' *Econometrica* 86(5), 1819–1838

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

— (2020) ‘Productivity and misallocation in general equilibrium.’ *The Quarterly Journal of Economics* 135(1), 105–163

— (2022) ‘Supply and demand in disaggregated Keynesian economies with an application to the COVID-19 crisis.’ *American Economic Review* 112(5), 1397–1436

— (2024) ‘Networks, barriers, and trade.’ *Econometrica* 92(2), 505–541

Baqaee, David R., Emmanuel Farhi, and Kunal Sangani (2024) ‘The supply-side effects of monetary policy.’ *Journal of Political Economy* 132(4), 1065–1112

Barth, Marvin J. III, and Valerie A. Ramey (2002) ‘The cost channel of monetary transmission.’ In ‘NBER Macroeconomics Annual 2001, Volume 16’ NBER Chapters (National Bureau of Economic Research, Inc) pp. 199–256

Basu, Susanto (1995) ‘Intermediate goods and business cycles: implications for productivity and welfare.’ *American Economic Review* 85(3), 512–31

Bigio, Saki, and Jennifer La’O (2020) ‘Distortions in production networks.’ *Quarterly Journal of Economics* 135, 2187–2253

Bils, Mark, and Peter J. Klenow (2004) ‘Some evidence on the importance of sticky prices.’ *Journal of Political Economy* 112(5), 947–985

Bouakez, Hafedh, Omar Rachedi, and Santoro Emiliano (2018) ‘Sectoral heterogeneity, production networks, and the effects of government spending.’ Technical Report

Carvalho, Carlos, and Fernanda Nechio (2011) ‘Aggregation and the PPP Puzzle in a sticky-price model.’ *American Economic Review* 101(6), 2391–2424

Carvalho, Vasco M, and Alireza Tahbaz-Salehi (2019) ‘Production networks: A primer.’ *Annual Review of Economics* 11(1), 635–663

Carvalho, Vasco M, Makoto Nirei, Yukiko U Saito, and Alireza Tahbaz-Salehi (2021) ‘Supply chain disruptions: Evidence from the Great East Japan earthquake.’ *The Quarterly Journal of Economics* 136(2), 1255–1321

Correia, Isabel, Juan Pablo Nicolini, and Pedro Teles (2008) ‘Optimal fiscal and monetary policy: Equivalence results.’ *Journal of Political Economy* 116(1), 141–170

David, Joel M., and David Zeke (2021) ‘Risk-Taking, Capital Allocation and Optimal Monetary Policy.’ Working Paper Series WP-2021-01, Federal Reserve Bank of Chicago, January

Devereux, Michael B, Karine Gente, and Changhua Yu (2023) 'Production networks and international fiscal spillovers.' *The Economic Journal* 133(653), 1871–1900

Evans, Charles L. (1992) 'Productivity shocks and real business cycles.' *Journal of Monetary Economics* 29(2), 191–208

Flynn, Joel P., Christina Patterson, and John Sturm (2020) 'Fiscal policy in a networked economy.' Working Paper, University of Chicago, August

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

Galí, Jordi (2015) *Monetary policy, inflation, and the business cycle: An introduction to the new keynesian framework and its applications second edition* number 10495. In 'Economics Books.' (Princeton University Press)

Ghassibe, Mishel (2021a) 'Endogenous production networks and non-linear monetary transmission.' Technical Report, Working paper

_ (2021b) 'Monetary policy and production networks: An empirical investigation.' *Journal of Monetary Economics* 119(C), 21–39

Giovanni, Julian Di, and Galina Hale (2022) 'Stock market spillovers via the global production network: Transmission of U.S. monetary policy.' *Journal of Finance* 77(6), 3373–3421

Hsieh, Chang-Tai, and Peter J Klenow (2009) 'Misallocation and manufacturing tfp in china and india.' *The Quarterly journal of economics* 124(4), 1403–1448

Hulten, Charles R. (1978) 'Growth accounting with intermediate inputs.' *Review of Economic Studies* 45(3), 511–518

Jones, Charles I. (2011) 'Intermediate goods and weak links in the theory of economic development.' *American Economic Journal: Macroeconomics* 3(2), 1–28

Kimball, Miles S (1995) 'The quantitative analytics of the basic neomonetarist model.' *Journal of Money, Credit and Banking* 27(4), 1241–1277

La'O, Jennifer, and Alireza Tahbaz-Salehi (2022) 'Optimal monetary policy in production networks.' *Econometrica* 90(3), 1295–1336

Levchenko, Andrei, Zhen Huo, and Nitya Pandalai-Nayar (2019) 'International comovement in the global production network.' CEPR Discussion Papers 13796, C.E.P.R. Discussion Papers, June

Liu, Ernest (2019) 'Industrial policies in production networks.' *The Quarterly Journal of Economics* 134(4), 1883–1948

Long, John B., and Charles I. Plosser (1983) 'Real business cycles.' *Journal of Political Economy* 91(1), 39–69

Luo, Shaowen (2020) 'Propagation of financial shocks in an input-output economy with trade and financial linkages of firms.' *Review of Economic Dynamics* 36, 246–269

Luo, Shaowen, and Daniel Villar (2023) 'Propagation of shocks in an input-output economy: Evidence from disaggregated prices.' *Journal of Monetary Economics* 137(C), 26–46

McKay, Alisdair, Emi Nakamura, and Jón Steinsson (2016) 'The power of forward guidance revisited.' *American Economic Review* 106(10), 3133–3158

Meier, Matthias, and Timo Reinelt (2022) 'Monetary policy, markup dispersion, and aggregate tfp.' *Review of Economics and Statistics* pp. 1–45

Miranda-Pinto, Jorge, and Eric R. Young (2022) 'Flexibility and frictions in multisector models.' *American Economic Journal: Macroeconomics* 14(3), 450–480

Nakamura, Emi, and Jon Steinsson (2010) 'Monetary non-neutrality in a multisector menu cost model.' *The Quarterly Journal of Economics* 125(3), 961–1013

Nakamura, Emi, and Jón Steinsson (2008) 'Five facts about prices: A reevaluation of menu cost models.' *The Quarterly Journal of Economics* 123(4), 1415–1464

Osotimehin, Sophie, and Latchezar Popov (2023) 'Misallocation and intersectoral linkages.' *Review of Economic Dynamics* 51, 177–198

Pasten, Ernesto, Raphael Schoenle, and Michael Weber (2017) 'Price rigidity and the origins of aggregate fluctuations.' NBER Working Papers 23750, National Bureau of Economic Research, Inc, August

Pasten, Ernesto, Raphael Schoenle, and Michael Weber (2020) 'The propagation of monetary policy shocks in a heterogeneous production economy.' *Journal of Monetary Economics* 116(C), 1–22

Pellet, Thomas, and Alireza Tahbaz-Salehi (2023) 'Rigid production networks.' *Journal of Monetary Economics* 137(C), 86–102

Ravenna, Federico, and Carl E. Walsh (2006) 'Optimal monetary policy with the cost channel.' *Journal of Monetary Economics* 53(2), 199–216

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

Not-For-Publication Technical Appendix

A Additional Figures and Tables

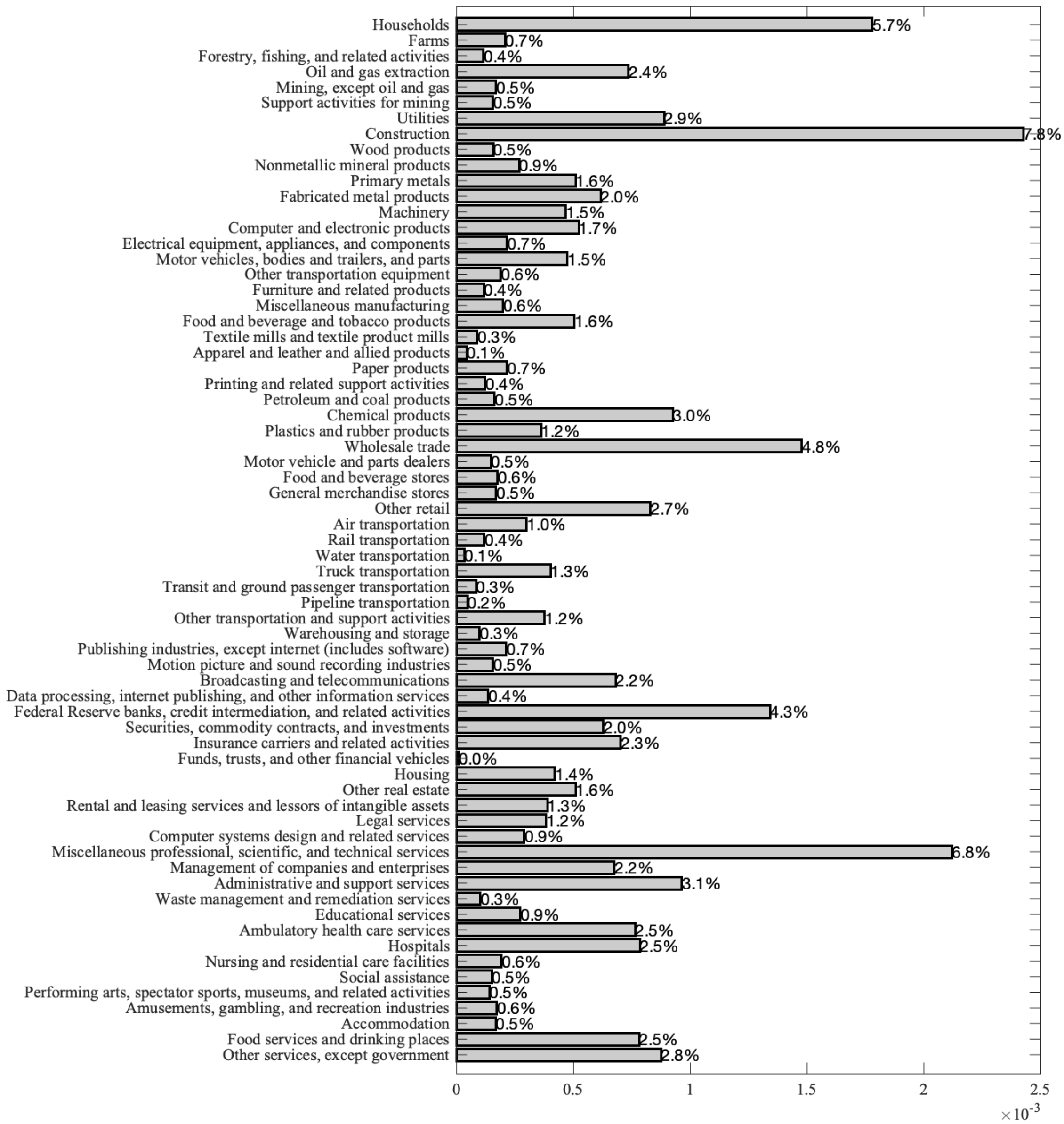

Figure A.1: Industry Contribution to Supply-Side Effect in 2006. Notes: The horizontal axis represents the magnitude of the supply-side effect. The percentages on each bar indicate the relative contribution of each sector to the total supply-side effect.

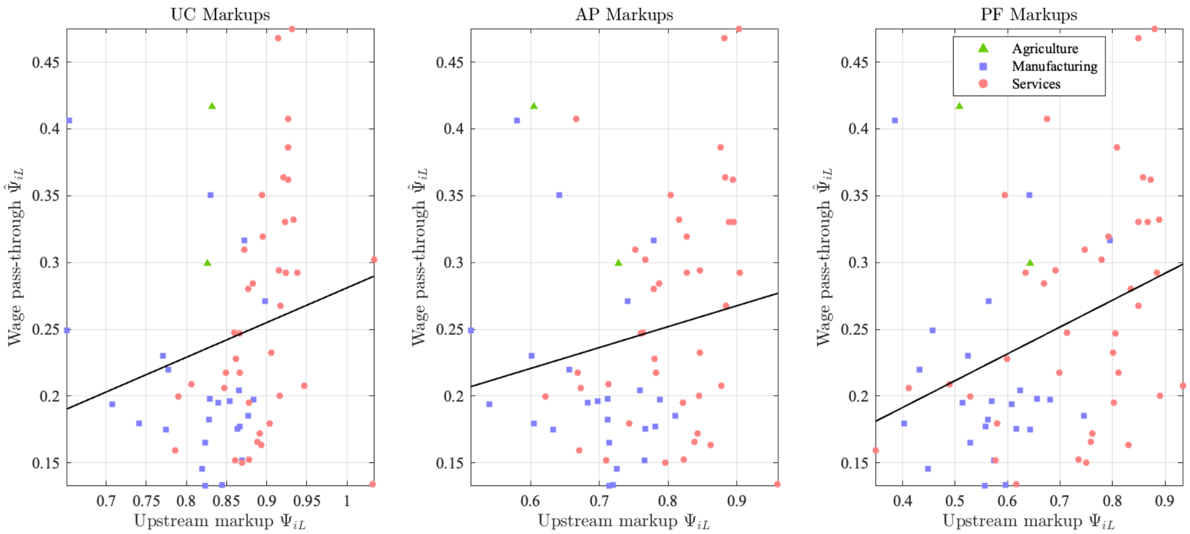

Figure A.2: Scatter Plot of Wage Pass-throughs and Upstream Markups in 2006. Notes: The sector corresponding to labor is not included in the figure. The wage pass-through and upstream markup corresponding to the labor sector are $\hat{\Psi}_{LL} = 1$ and $\Psi_{LL} = 1$, respectively.

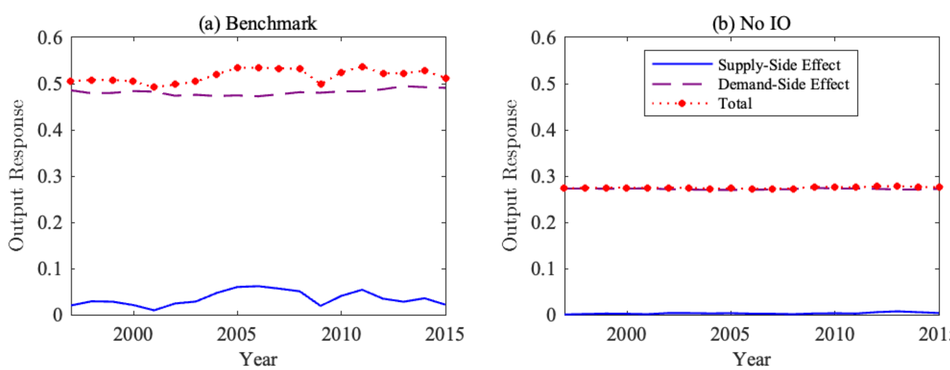

Figure A.3: The Supply-Side and Demand-Side Effects of Monetary Policy under UC Markups. Notes: This figure reports numerical values for the supply-side and demand-side effects of monetary policy. The solid blue line corresponds to the supply-side effect, the dashed purple line signifies the demand-side effect, and the dotted red line with circle markers indicates the total effect.

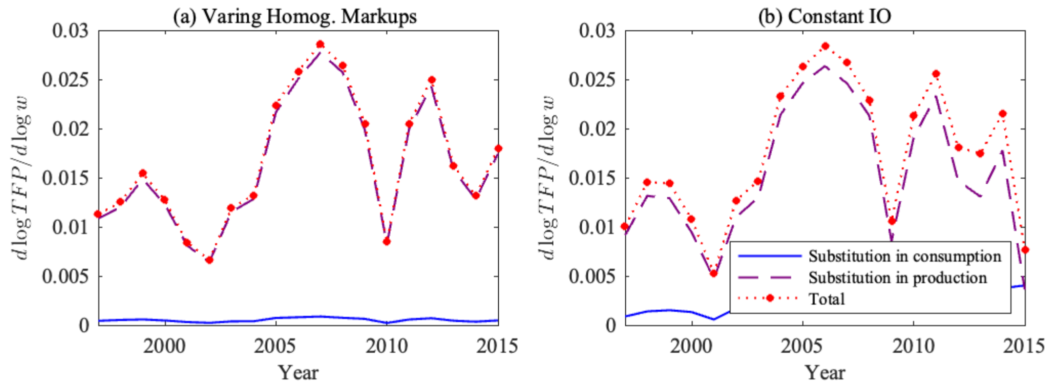

Figure A.4: Decomposition of Supply-Side Effect with UC Markups: Homogeneous Markups (Panel a) and Constant Input Output (IO) Table (Panel b).

	Optimal Policy	Output-Gap Targeting	CPI Targeting	DC Targeting
Welfare Gain (Quad. Approx., % of GDP)	-0.565	-0.663	-0.671	-0.664
First order welfare gain				
Aggregate wedge	0.079	10^{-5}	0.001	0
Supply-side effect	0.119	10^{-4}	0.001	10^{-4}
Second order welfare gain				
Output gap volatility	-0.011	0	-0.002	-10^{-5}
Within-sector price dispersion	-0.609	-0.542	-0.542	-0.543
Cross-sector price dispersion	-0.146	-0.126	-0.134	-0.126
Variation in allocative efficiency	0	0	0	0
$(1-\Lambda_L)\times$ second-order terms	0.004	0.004	0.005	0.004
Welfare Gain (Exact Model, % of GDP)	-0.517	-0.568	-0.582	-0.569
Cosine similarity to optimal policy	1	0.989	0.617	0.987

Table A.1: Welfare Gains under Various Monetary Policies in a Cobb-Douglas Economy with UC Markups. Notes: This table reports welfare gains under various monetary policies as a percentage of steady-state consumption, based on 10,000 draws. The row labeled 'Exact Model' calculates welfare gain based on the exact social welfare function. The remaining rows show welfare gains and their decomposition, derived from a second-order quadratic approximation of the welfare function. The last column reports the welfare gain under the divine coincidence (DC) targeting policy which closes the actual output gap using the divine coincidence index as inflation index.

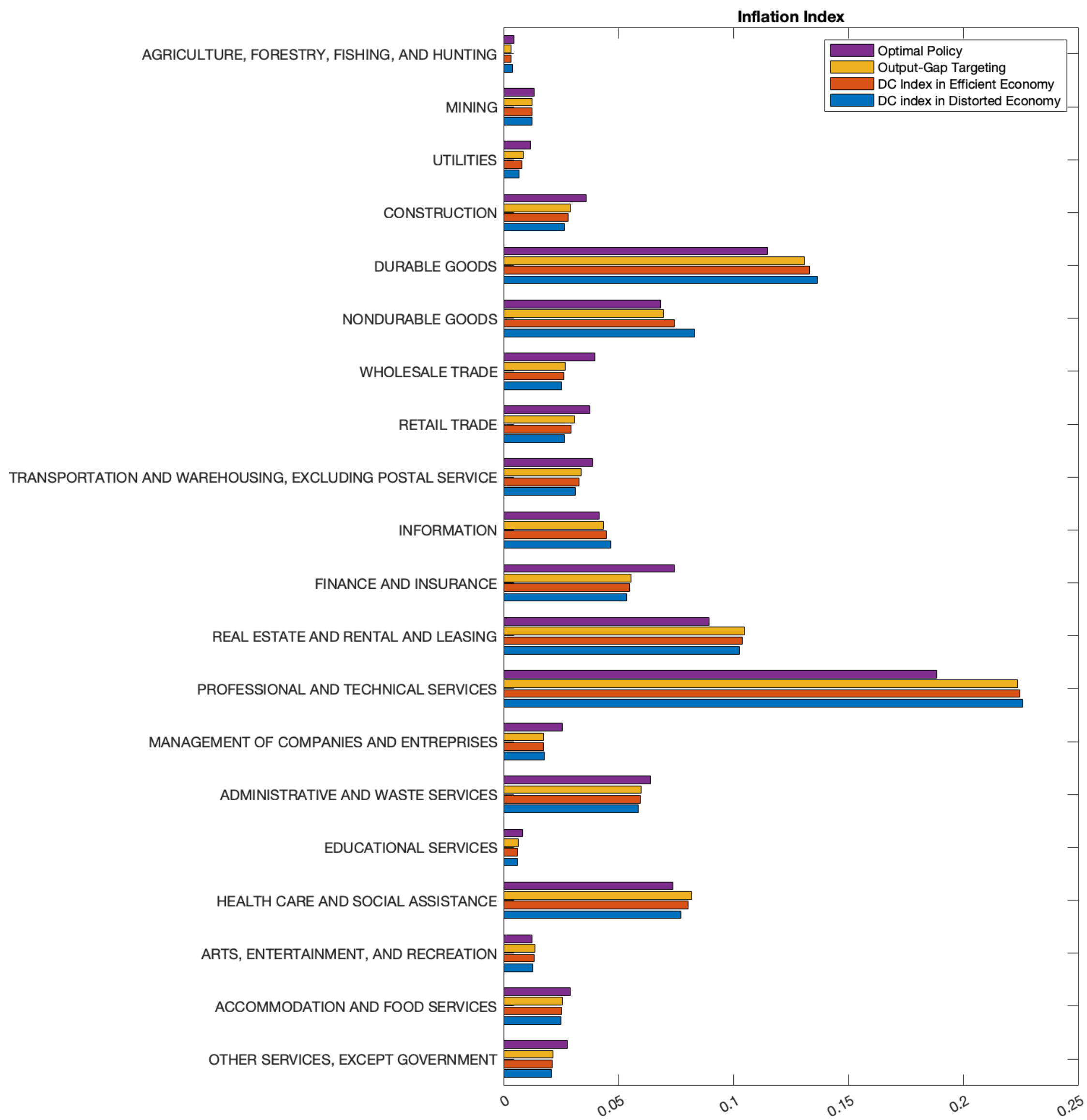

Figure A.5: Industry Weights of Inflation Index under Various Policies. Notes: The 66 sectors are aggregated into 20 broader categories. Compared to the 20 broad sectors in NAICS, our classification splits Manufacturing into Durable and Nondurable Goods and excludes Public Administration. The optimal policy's inflation index follows equation (10). The industry weight for output-gap targeting is $(1/\theta_i - 1)\tilde{\lambda}_i - \mathcal{J}_i$. For the divine coincidence (DC) index, the weight in an efficient economy is $(1/\theta_i - 1)\tilde{\lambda}_i$, while in a distorted economy, it is $(1/\theta_i - 1)\tilde{\lambda}_i + \varphi\mathcal{J}_i$.

B Related Proofs

Throughout the appendix, we interchangeably use Ω^f and $\Omega^n_{(L)}$ to represent the $N \times 1$ vector of labor income share, while the Ψ^f and $\Psi^n_{(L)}$ are used interchangeably to represent the $N \times 1$ vector of Leontief inverse of labor.

Lemma 2 (Property of the Leontief Inverse Matrix). *In general, for any input-output matrix Ω satisfying $\sum_{j=1}^{N+1} \Omega_{ij} \leq 1$ for all i, and its corresponding Leontief inverse matrix Ψ which is defined as $\Psi = (I - \Omega)^{-1}$, we have (i) $\Psi^n = (I - \Omega^n)^{-1}$, (ii) $\Psi^n \Omega^f = \Psi^f$.*

Proof of Lemma 2. By definition of the Leontief inverse matrix, $\Psi\Omega = \Psi - I$. This can be rewritten in a block matrix form as

$$\Psi\Omega = \begin{bmatrix} \Psi^n & \Psi^f \\ \mathbf{0} & I \end{bmatrix} \begin{bmatrix} \Omega^n & \Omega^f \\ \mathbf{0} & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \Psi^n \Omega^n & \Psi^n \Omega^f \\ \mathbf{0} & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \Psi^n - I & \Psi^f \\ \mathbf{0} & \mathbf{0} \end{bmatrix} = \Psi - I.$$

Thus, $\Psi^n \Omega^n = \Psi^n - I$ and $\Psi^n \Omega^f = \Psi^f$.

□

Lemma 3. *Given price adjustment probability $\theta_i \in [0, 1]$ for all i, we observe that: (i) the rigidity-adjusted Domar weight $\hat{\lambda}_k$ is less than corresponding cost-based Domar weight $\tilde{\lambda}_k$; and (ii) the wage pass-through $\hat{\Psi}_{iL}$ is less than its corresponding price adjustment probability θ_i.*

Proof of Lemma 3. By definition of Leontief inverse, we have

$$\hat{\Psi}^n = (I - \Theta\tilde{\Omega}^n)^{-1} = I + \Theta\tilde{\Omega}^n + (\Theta\tilde{\Omega}^n)^2 + ... \leq I + \tilde{\Omega}^n + (\tilde{\Omega}^n)^2 + ... = (I - \tilde{\Omega}^n)^{-1} = \tilde{\Psi}^n. \quad \text{(B-1)}$$

Hence,

$$\hat{\lambda}' = b'\hat{\Psi}^n \leq b'\tilde{\Psi}^n = \tilde{\lambda}'. \quad \text{(B-2)}$$

Analogously, we find that $(I - \tilde{\Omega}^n\Theta)^{-1} \leq (I - \tilde{\Omega}^n)^{-1}$, which implies that

$$\begin{aligned} \hat{\Psi}^f - \theta &= \hat{\Psi}^f - \Theta\tilde{\Psi}^f && \text{(B-3)} \\ &= (I - \Theta\tilde{\Omega}^n)^{-1}\Theta\tilde{\Omega}^f - \Theta(I - \tilde{\Omega}^n)^{-1}\tilde{\Omega}^f && \text{(B-4)} \\ &= \Theta[(I - \tilde{\Omega}^n\Theta)^{-1} - (I - \tilde{\Omega}^n)^{-1}]\tilde{\Omega}^f \leq \mathbf{0} && \text{(B-5)} \end{aligned}$$

where the first equality uses the fact that $\tilde{\Psi}^f = \mathbf{1}$.

□

Lemma 4. *The pass-through of nominal wages into sector prices $\hat{\Psi}_{(L)}$ is weakly increasing in price adjustment probability $\{\theta_i\}_{i=1}^N$. Moreover, the pass-through of nominal wages into consumer price $\hat{\Lambda}_L$ is weakly increasing in price adjustment probability $\{\theta_i\}_{i=1}^N$ and it is bounded between 0 and $E_b(\theta)$.*

Proof of Lemma 4. Since $\hat{\Psi}^f = [I - \Theta\tilde{\Omega}^n]^{-1}\Theta\tilde{\Omega}^f = [\Theta^{-1} - \tilde{\Omega}^n]^{-1}\tilde{\Omega}^f$, we have

$$\begin{aligned} \mathrm{d}\hat{\Psi}^f &= \mathrm{d}[\Theta^{-1} - \tilde{\Omega}^n]^{-1}\tilde{\Omega}^f && \text{(B-6)} \\ &= [\Theta^{-1} - \tilde{\Omega}^n]^{-1}\Theta^{-1}(\mathrm{d}\Theta)\Theta^{-1}[\Theta^{-1} - \tilde{\Omega}^n]^{-1}\tilde{\Omega}^f && \text{(B-7)} \\ &= \underbrace{[I - \Theta\tilde{\Omega}^n]^{-1}}_{\hat{\Psi}^n}(\mathrm{d}\Theta)\Theta^{-1}\underbrace{[I - \Theta\tilde{\Omega}^n]^{-1}\Theta\tilde{\Omega}^f}_{\hat{\Psi}^f} && \text{(B-8)} \\ &= \hat{\Psi}^n(\mathrm{d}\Theta)\Theta^{-1}\hat{\Psi}^f && \text{(B-9)} \end{aligned}$$

and

$$\mathrm{d}\hat{\Lambda}_L = b'\mathrm{d}\hat{\Psi}^f = b'\hat{\Psi}^n(\mathrm{d}\Theta)\Theta^{-1}\hat{\Psi}^f = \hat{\lambda}'(\mathrm{d}\Theta)\Theta^{-1}\hat{\Psi}^f. \tag{B-10}$$

Or equivalently,

$$\mathrm{d}\hat{\Psi}_{iL} = \sum_k \hat{\Psi}_{ik}\hat{\Psi}_{kL}\mathrm{d}\log\theta_k \quad \text{and} \quad \mathrm{d}\hat{\Lambda}_L = \sum_k \hat{\lambda}_k\hat{\Psi}_{kL}\mathrm{d}\log\theta_k. \tag{B-11}$$

Thus, $\hat{\Psi}_{iL}$ and $\hat{\Lambda}_L$ are weakly increasing in $\{\theta_i\}_{i=1}^N$.

We then show that $\hat{\Lambda}_L$ is bounded above by $\mathbb{E}_b(\theta)$. Since $\hat{\Psi}_{iL} \leq \theta_i$ (Lemma 3), we have

$$\hat{\Lambda}_L = b'\hat{\Psi}^f \leq b'\theta = \mathrm{E}_b(\theta) \leq 1. \tag{B-12}$$

This result states that as long as a sector takes use of an intermediate input produced by a sector (either its own sector or another sector) with sticky price, the wage pass-through into consumer price is incomplete.

□

Lemma 5. *When evaluated at an inefficient initial equilibrium where all initial markups are positive ($\bar{\mu} \geq 1$), both Ψ_{iL} and Λ_L are bounded between 0 and 1. They weakly decrease with sectoral initial markups $\{\bar{\mu}_i\}_{i=1}^N$, indicating that a lower Ψ_{iL} corresponds to a higher degree of markup in the supply chain of sector i. Additionally, for each sector, the upstream markup Ψ_{iL} is less than or equal to its counterpart in an economy without input-output linkages, $\bar{\mu}_i^{-1}$.*

Proof of Lemma 5. Analogy to Lemma 3 and Lemma 4 by replacing θ_i with $\bar{\mu}_i^{-1}$. □

C Nested CES Economies

C.1 Standard-Form for Nested CES Economies

In this section, we expand our input-output notation to incorporate producer 0 (households). We begin by adjusting the cost-based input-output matrix, setting indices of $\tilde{\Omega}$ to start at 0. Specifically, $\tilde{\Omega}_{i0} = 0$ for all $0 \le i \le N+1$ and $\tilde{\Omega}_{0j} = b_j$ for all $1 \le j \le N+1$. Correspondingly, $x_{i0} = 0$ for all $0 \le i \le N+1$ and $x_{0i} = c_i$ for all $1 \le j \le N+1$. We then set the household-related parameters: $\bar{\mu}_C = \theta_C = 1$ and expand all Leontief inverse matrices to dimensions of $(1+N+1) \times (1+N+1)$. Finally, we extend the $(N+1) \times 1$ vector $\mathrm{d}\log p$ to an $(1+N+1) \times 1$ vector by defining $p_0 = P^Y$.

The production function of goods in sector k is given by

$$\frac{y_k}{\bar{y}_k} = \bar{A}_k \left(\sum_l \omega_{kl}^{\frac{1}{\sigma_k}} \left(\frac{x_{kl}}{\bar{x}_{kl}} \right)^{\frac{\sigma_k - 1}{\sigma_k}} \right)^{\frac{\sigma_k}{\sigma_k - 1}}$$

where $\bar{A}_k = q_k A_k$, with sectoral productivity loss q_k defined by equation (E-29). Note that $\mathrm{d}\log \bar{A}_i = \mathrm{d}\log A_i$ since $\mathrm{d}\log q_i = 0$.[1] Additionally, each sector is associated with a unique elasticity of substitution σ_k.

The optimal condition yields

$$\mathrm{d}\log x_{ki} - \mathrm{d}\log x_{kj} = -\sigma_k (\mathrm{d}\log p_i - \mathrm{d}\log p_j) \ \forall i, j, \tag{C-1}$$

which indicates how resources are reallocated due to relative price changes.

In addition, the production function of producer 0 (households) is given by a final-demand aggregator

$$\frac{Y}{\bar{Y}} = \left(\sum_l \omega_{0l}^{\frac{1}{\sigma_C}} \left(\frac{c_l}{\bar{c}_l} \right)^{\frac{\sigma_C - 1}{\sigma_C}} \right)^{\frac{\sigma_C}{\sigma_C - 1}},$$

where the elasticity of substitution is σ_C.

C.2 Related Proofs in An Arbitrary CES Economy

In this section, we apply the methodology of Baqaee and Farhi (2020) to derive the expression for labor income share. This expression comprises two components: first, the

[1] See more discussions in Appendix E.3.

direct effect of changes in ex-post markups, holding constant distribution of resources (input shares); and second, the equilibrium changes in the distribution of resources. Accordingly, our equations from (C-2) to (C-18) adapt their findings to the context of a single factor case.

Proof of Theorem 1. Since $\Omega_{ji} = \frac{p_i x_{ji}}{\mu_j \sum_l p_l x_{jl}}$ for all $0 \le i, j \le N+1$, we have

$$\mathrm{d}\log\Omega_{ji} = -\mathrm{d}\log\mu_j + \mathrm{d}\log p_i + \mathrm{d}\log x_{ji} - \sum_l \tilde{\Omega}_{jl}\mathrm{d}\log(p_l x_{jl}) \tag{C-2}$$

$$= -\mathrm{d}\log\mu_j + \sum_l \tilde{\Omega}_{jl}\left(\mathrm{d}\log p_i - \mathrm{d}\log p_l\right) - \sum_l \tilde{\Omega}_{jl}\left(\mathrm{d}\log x_{ji} - \mathrm{d}\log x_{jl}\right) \tag{C-3}$$

$$= -\mathrm{d}\log\mu_j + (1-\sigma_j)\sum_l \tilde{\Omega}_{jl}\left(\mathrm{d}\log p_i - \mathrm{d}\log p_l\right) \tag{C-4}$$

$$= -\mathrm{d}\log\mu_j + (1-\sigma_j)(\mathrm{d}\log p_i - \sum_l \tilde{\Omega}_{jl}\mathrm{d}\log p_l). \tag{C-5}$$

Considering the covariance with $\tilde{\Omega}(j,:)$ as weights, we have

$$\mathrm{Cov}_{\tilde{\Omega}(j,:)}(\mathrm{d}\log p, I_{(i)}) = \tilde{\Omega}_{ji}\mathrm{d}\log p_i - \sum_l \tilde{\Omega}_{jl}\mathrm{d}\log p_l \cdot \tilde{\Omega}_{ji} \tag{C-6}$$

$$= \tilde{\Omega}_{ji}(\mathrm{d}\log p_i - \sum_l \tilde{\Omega}_{jl}\mathrm{d}\log p_l), \tag{C-7}$$

where $\tilde{\Omega}(j,:)$ is the j-th row of $\tilde{\Omega}$ and $I_{(i)}$ is the i-th column of the identity matrix I.

Thus, the total differential of Ω_{ji} is given by

$$d\Omega_{ji} = \Omega_{ji}\left[-\mathrm{d}\log\mu_j + (1-\sigma_j)(\mathrm{d}\log p_i - \sum_l \tilde{\Omega}_{jl}\mathrm{d}\log p_l)\right] \tag{C-8}$$

$$= -\Omega_{ji}\mathrm{d}\log\mu_j + \frac{\Omega_{ji}}{\tilde{\Omega}_{ji}}(1-\sigma_j)\mathrm{Cov}_{\tilde{\Omega}(j,:)}(\mathrm{d}\log p, I_{(i)}) \tag{C-9}$$

$$= -\Omega_{ji}\mathrm{d}\log\mu_j + \frac{1}{\bar{\mu}_j}(1-\sigma_j)\mathrm{Cov}_{\tilde{\Omega}(j,:)}(\mathrm{d}\log p, I_{(i)}). \tag{C-10}$$

Using $\Psi = (I-\Omega)^{-1}$, we obtain

$$\mathrm{d}\Psi = \Psi\mathrm{d}\Omega\Psi. \tag{C-11}$$

Or equivalently,

$$d\Psi_{mn} = \sum_j \sum_i \Psi_{mj} d\Omega_{ji} \Psi_{in} \tag{C-12}$$

$$= -\sum_j \sum_i \Psi_{mj}\Psi_{in}\Omega_{ji} \mathrm{d}\log\mu_j + \sum_j \sum_i \Psi_{mj}\Psi_{in}\bar{\mu}_j^{-1}(1-\sigma_j)\mathrm{Cov}_{\tilde{\Omega}(j,:)}(\mathrm{d}\log p, I_{(i)}) \tag{C-13}$$

$$= -\sum_j \Psi_{mj} \mathrm{d}\log\mu_j \sum_i \Omega_{ji}\Psi_{in} + \sum_j \Psi_{mj}\bar{\mu}_j^{-1}(1-\sigma_j)\mathrm{Cov}_{\tilde{\Omega}(j,:)}(\mathrm{d}\log p, \Psi_{(n)}). \tag{C-14}$$

Using $\Omega\Psi = \Psi - I$, we can rewrite the expression above as

$$d\Psi_{mn} = -\sum_j \Psi_{mj}(\Psi_{jn} - \delta_{jn})\mathrm{d}\log\mu_j + \sum_j \Psi_{mj}\bar{\mu}_j^{-1}(1-\sigma_j)\mathrm{Cov}_{\tilde{\Omega}(j,:)}(\mathrm{d}\log p, \Psi_{(n)}), \tag{C-15}$$

where δ_{jn} is the jn-th element of the identity matrix.

Given $b'\Psi = \lambda$, we get

$$d\lambda_n = \sum_k b_k d\Psi_{kn} = -\sum_j \lambda_j(\Psi_{jn} - \delta_{jn})\mathrm{d}\log\mu_j + \sum_j \lambda_j\bar{\mu}_j^{-1}(1-\sigma_j)\mathrm{Cov}_{\tilde{\Omega}(j,:)}(\mathrm{d}\log p, \Psi_{(n)}). \tag{C-16}$$

Dividing both sides by λ_n, we have

$$\mathrm{d}\log\lambda_n = -\sum_j \frac{\lambda_j}{\lambda_n}(\Psi_{jn} - \delta_{jn})\mathrm{d}\log\mu_j + \sum_j \frac{\lambda_j}{\lambda_n}\bar{\mu}_j^{-1}(1-\sigma_j)\mathrm{Cov}_{\tilde{\Omega}(j,:)}(\mathrm{d}\log p, \Psi_{(n)}). \tag{C-17}$$

Accordingly,

$$\mathrm{d}\log\Lambda_L = -\sum_j \lambda_j \frac{\Psi_{jL}}{\Lambda_L}\mathrm{d}\log\mu_j + \frac{1}{\Lambda_L}\sum_j \frac{\lambda_j}{\bar{\mu}_j}(1-\sigma_j)\mathrm{Cov}_{\tilde{\Omega}(j,:)}(\mathrm{d}\log p, \Psi_{(L)}). \tag{C-18}$$

To further simplify the expression above, note that

$$\begin{aligned}
&\frac{1}{\Lambda_L}\sum_{j=1}^{N}\frac{\lambda_j}{\bar{\mu}_j}\mathrm{Cov}_{\tilde{\Omega}(j,:)}\left(\mathrm{d}\log p, \Psi_{(L)}\right)\\
=&\frac{1}{\Lambda_L}\sum_{j=1}^{N}\frac{\lambda_j}{\bar{\mu}_j}\left[\mathrm{E}_{\tilde{\Omega}(j,:)}(\Psi_{iL}\mathrm{d}\log p_i) - \mathrm{E}_{\tilde{\Omega}(j,:)}(\Psi_{iL})\mathrm{E}_{\tilde{\Omega}(j,:)}(\mathrm{d}\log p_i)\right] && \text{(C-19)}\\
=&\frac{1}{\Lambda_L}\sum_{j=1}^{N}\frac{\lambda_j}{\bar{\mu}_j}\left[\sum_{k=1}^{N+1}\tilde{\Omega}_{jk}\Psi_{kL}\mathrm{d}\log p_k - \underbrace{\left(\sum_{k=1}^{N+1}\tilde{\Omega}_{jk}\Psi_{kL}\right)}_{\bar{\mu}_j\Psi_{jL}}\underbrace{\left(\sum_{k=1}^{N+1}\tilde{\Omega}_{jk}\mathrm{d}\log p_k\right)}_{\mathrm{d}\log p_j+\mathrm{d}\log A_j-\mathrm{d}\log\mu_j}\right] && \text{(C-20)}\\
=&\frac{1}{\Lambda_L}\sum_{j=1}^{N}\frac{\lambda_j}{\bar{\mu}_j}\left[\sum_{k=1}^{N+1}\tilde{\Omega}_{jk}\Psi_{kL}\mathrm{d}\log p_k - \bar{\mu}_j\Psi_{jL}(\mathrm{d}\log p_j + \mathrm{d}\log A_j - \mathrm{d}\log\mu_j)\right] && \text{(C-21)}\\
=&\sum_{k=1}^{N+1}\frac{\Psi_{kL}}{\Lambda_L}\underbrace{\sum_{j=1}^{N}\lambda_j\frac{\tilde{\Omega}_{jk}}{\bar{\mu}_j}}_{\lambda_k-b_k}\mathrm{d}\log p_k - \sum_{j=1}^{N}\lambda_j\frac{\Psi_{jL}}{\Lambda_L}(\mathrm{d}\log p_j + \mathrm{d}\log A_j - \mathrm{d}\log\mu_j) && \text{(C-22)}\\
=&\mathrm{d}\log w + \sum_{j=1}^{N}(\lambda_j - b_j)\frac{\Psi_{jL}}{\Lambda_L}\mathrm{d}\log p_j - \sum_{j=1}^{N}\lambda_j\frac{\Psi_{jL}}{\Lambda_L}(\mathrm{d}\log p_j + \mathrm{d}\log A_j - \mathrm{d}\log\mu_j) && \text{(C-23)}\\
=&\mathrm{d}\log w - \sum_{j=1}^{N}b_j\frac{\Psi_{jL}}{\Lambda_L}\mathrm{d}\log p_j + \sum_{j=1}^{N}\lambda_j\frac{\Psi_{jL}}{\Lambda_L}(\mathrm{d}\log\mu_j - \mathrm{d}\log A_j). && \text{(C-24)}
\end{aligned}$$

where the third equality is derived by using Sheppard's Lemma, where $\mathrm{d}\log p_j = -\mathrm{d}\log A_j + \mathrm{d}\log\mu_j + \sum_{k=1}^{N+1}\tilde{\Omega}_{jk}\mathrm{d}\log p_k$, and the fifth equality uses the facts that $\sum_{j=1}^{N}\lambda_j\frac{\tilde{\Omega}_{jk}}{\bar{\mu}_j} = \sum_{j=1}^{N}\lambda_j\Omega_{jk} = \lambda_k - b_k$ and $\frac{\Psi_{LL}}{\Lambda_L}(\Lambda_L - b_L)\mathrm{d}\log p_{N+1} = \mathrm{d}\log w$.

In addition,

$$\begin{aligned}
\frac{1}{\Lambda_L}\frac{\lambda_0}{\bar{\mu}_0}\mathrm{Cov}_{\tilde{\Omega}(0,:)}\left(\mathrm{d}\log p, \Psi_{(L)}\right) =& \frac{1}{\Lambda_L}\left[E_b(\Psi_{iL}\mathrm{d}\log p_i) - \underbrace{E_b(\Psi_{iL})}_{\Lambda_L}\underbrace{E_b(\mathrm{d}\log p_i)}_{\mathrm{d}\log P^Y}\right] && \text{(C-25)}\\
=&\sum_{j=1}^{N}b_j\frac{\Psi_{jL}}{\Lambda_L}\mathrm{d}\log p_j - \mathrm{d}\log P^Y. && \text{(C-26)}
\end{aligned}$$

Combining the expressions above, we get:

$$\sum_{j=0}^{N} \frac{\lambda_j}{\Lambda_L} \frac{1}{\bar{\mu}_j} \mathrm{Cov}_{\tilde{\Omega}(j,:)} \left(\mathrm{d}\log p, \Psi_{(L)} \right) = \mathrm{d}\log w - \mathrm{d}\log P^Y + \sum_{j=1}^{N} \lambda_j \frac{\Psi_{jL}}{\Lambda_L} (\mathrm{d}\log \mu_j - \mathrm{d}\log A_j). \quad \text{(C-27)}$$

Thus, we can express $\mathrm{d}\log \Lambda_L$ as:

$$\mathrm{d}\log \Lambda_L = \mathrm{d}\log w - \mathrm{d}\log P^Y - \sum_{j=1}^{N} \lambda_j \frac{\Psi_{jL}}{\Lambda_L} \mathrm{d}\log A_j - \sum_j \frac{\sigma_j \lambda_j}{\bar{\mu}_j} \mathrm{Cov}_{\tilde{\Omega}(j,:)} (\mathrm{d}\log p_i, \frac{\Psi_{iL}}{\Lambda_L}) \quad \text{(C-28)}$$

By the definition of the labor income share, we have:

$$\mathrm{d}\log \Lambda_L = \mathrm{d}\log w + \mathrm{d}\log L - (\mathrm{d}\log P^Y + \mathrm{d}\log Y). \quad \text{(C-29)}$$

Combining equations (C-28) and (C-29) yields:

$$\begin{aligned} \mathrm{d}\log \mathrm{TFP} &= \mathrm{d}\log Y - \mathrm{d}\log L && \text{(C-30)} \\ &= \mathrm{d}\log w - \mathrm{d}\log P^Y - \mathrm{d}\log \Lambda_L && \text{(C-31)} \\ &= \underbrace{\sum_{j=1}^{N} \lambda_j \frac{\Psi_{jL}}{\Lambda_L} \mathrm{d}\log A_j}_{\text{Direct technology channel}} + \underbrace{\sum_j \frac{\sigma_j \lambda_j}{\bar{\mu}_j} \mathrm{Cov}_{\tilde{\Omega}(j,:)} (\mathrm{d}\log p_i, \frac{\Psi_{iL}}{\Lambda_L})}_{\text{Misallocation channel}} && \text{(C-32)} \end{aligned}$$

□

Proof of Proposition 1. By applying the Sheppard's lemma, we obtain the following relationship:

$$\mathrm{d}\log mc = -\mathrm{d}\log A + \tilde{\Omega}^n \pi + \tilde{\Omega}^f \mathrm{d}\log w. \quad \text{(C-33)}$$

Combining this with the equation $\pi = \Theta \mathrm{d}\log mc$, we derive:

$$\begin{aligned} \pi &= -(I - \Theta\tilde{\Omega}^n)^{-1}\Theta(\mathrm{d}\log A - \tilde{\Omega}^f \mathrm{d}\log w) && \text{(C-34)} \\ &= -\hat{\Psi}^n \Theta \mathrm{d}\log A + \hat{\Psi}^f \mathrm{d}\log w. && \text{(C-35)} \end{aligned}$$

As for the consumer price, it evolves as follows:

$$
\begin{aligned}
\mathrm{d}\log P^Y &= b'\pi && \text{(C-36)}\\
&= -\underbrace{b'\hat{\Psi}^n}_{\hat{\lambda}'}\Theta \mathrm{d}\log A + \underbrace{b'\hat{\Psi}^f}_{\hat{\Lambda}_L}\mathrm{d}\log w && \text{(C-37)}\\
&= -\hat{\lambda}'\Theta \mathrm{d}\log A + \hat{\Lambda}_L \mathrm{d}\log w. && \text{(C-38)}
\end{aligned}
$$

By Theorem 1, we have

$$
\frac{\mathrm{d}\log \mathrm{TFP}}{\mathrm{d}\log w} = \sum_j \frac{\sigma_j\lambda_j}{\bar{\mu}_j}\mathrm{Cov}_{\tilde{\Omega}(j,:)}\left(\underbrace{\frac{\mathrm{d}\log p_i}{\mathrm{d}\log w}}_{\hat{\Psi}_{iL}}, \frac{\Psi_{iL}}{\Lambda_L}\right) = \sum_{j=0}^{N}\frac{\sigma_j\lambda_j}{\bar{\mu}_j}\mathrm{Cov}_{\tilde{\Omega}(j,:)}\left(\hat{\Psi}_{iL}, \frac{\Psi_{iL}}{\Lambda_L}\right) \tag{C-39}
$$

□

Analogously, the response of TFP to a productivity shock is

$$
\begin{aligned}
\frac{\mathrm{d}\log \mathrm{TFP}}{\mathrm{d}\log A_k} &= \lambda_k\frac{\Psi_{kL}}{\Lambda_L} + \sum_j \frac{\sigma_j\lambda_j}{\bar{\mu}_j}\mathrm{Cov}_{\tilde{\Omega}(j,:)}\left(\underbrace{\frac{\mathrm{d}\log p_i}{\mathrm{d}\log A_k}}_{-\theta_k\hat{\Psi}_{ik}}, \frac{\Psi_{iL}}{\Lambda_L}\right) && \text{(C-40)}\\
&= \underbrace{\lambda_k\frac{\Psi_{kL}}{\Lambda_L}}_{\text{Direct technology channel}} \underbrace{- \theta_k\sum_{j=0}^{N}\frac{\sigma_j\lambda_j}{\bar{\mu}_j}\mathrm{Cov}_{\tilde{\Omega}(j,:)}\left(\hat{\Psi}_{ik}, \frac{\Psi_{iL}}{\Lambda_L}\right)}_{\text{Misallocation channel}} \doteq \omega_k. && \text{(C-41)}
\end{aligned}
$$

Note that $\tilde{\Omega}_{i0} = 0$ for all i. The formulas in this section remain valid when all input-output matrices and Leontief inverse matrices are replaced with $(N+1)\times(N+1)$ matrices, as defined in the main text.

D Money Supply and Nominal Wages

In this section, we demonstrate that the monetary authority can equivalently select either nominal wages or money supply as its policy instrument. We then leverage this one-to-one mapping between money supply and nominal wages to analyze the decomposition of the output response.

D.1 Isomorphic Relationship

Lemma 6. *Suppose the elasticity of labor income share to nominal wage is less than one, there exists a one-to-one mapping between money supply M and nominal wages w for all realizations of productivity shocks A.*

Proof of Lemma 6. The consumption-leisure trade-off is expressed as:

$$\mathrm{d}\log w - \mathrm{d}\log P^Y = \gamma \mathrm{d}\log Y + \varphi \mathrm{d}\log L \tag{D-1}$$

$$= (\gamma + \varphi)\mathrm{d}\log Y - \varphi \mathrm{d}\log \mathrm{TFP}. \tag{D-2}$$

When integrated with the cash-in-advance constraint, this results in:

$$(\gamma + \varphi)\mathrm{d}\log M = \mathrm{d}\log w + (\gamma + \varphi - 1)\mathrm{d}\log P^Y + \varphi \mathrm{d}\log \mathrm{TFP}. \tag{D-3}$$

Referencing the previous section:

$$\mathrm{d}\log \mathrm{TFP} = \varpi' \mathrm{d}\log A + (1 - \hat{\Lambda}_L - \xi)\mathrm{d}\log w. \tag{D-4}$$

Combining the above with equation (C-38) yields:

$$[1 - \hat{\Lambda}_L + (\gamma + \varphi)\hat{\Lambda}_L + \varphi(1 - \hat{\Lambda}_L - \xi)]\mathrm{d}\log w = (\gamma + \varphi)d\log M - [\varphi\varpi' - (\gamma + \varphi - 1)\hat{\lambda}'\Theta]\mathrm{d}\log A. \tag{D-5}$$

This derivation suggests that nominal wage decreases relative to money supply only when the supply-side effect is substantially negative. Given the the elasticity of labor income share to nominal wage is less than one, we derive the following inequality:

$$1 - \hat{\Lambda}_L + (\gamma + \varphi)\hat{\Lambda}_L + \varphi(1 - \hat{\Lambda}_L - \xi) \geq 1 - \hat{\Lambda}_L + \gamma\hat{\Lambda}_L > 1 - \hat{\Lambda}_L \geq 0 \tag{D-6}$$

Consequently, with $1 + (\gamma - 1)\hat{\Lambda}_L + \varphi(1 - \xi) > 0$, equation (D-5) confirms a one-to-one mapping between money supply M and nominal wages w for all realizations of productivity shocks A. □

D.2 Decomposition of Output Response

We then investigate how labor supply responds to monetary shocks. In response to a change in nominal wage, the consumption-leisure tradeoff implies

$$\mathrm{d}\log L = \underbrace{\frac{1-\hat{\Lambda}_L}{\varphi}\mathrm{d}\log w}_{\text{Substitution effect}} \quad - \underbrace{\frac{\gamma}{\varphi}\mathrm{d}\log Y}_{\text{Wealth effect}} \tag{D-7}$$

This equation decomposes the labor response into two components as in the literature. The first term represents the substitution effect, which arises when nominal wages increase ($\mathrm{d}\log w > 0$) and consumer prices do not completely offset this increase due to nominal rigidities (i.e. $\hat{\Lambda}_L < 1$), resulting in an increase in real wages that boosts up labor supply. The factor $1-\hat{\Lambda}_L$ measures the increase in real wages that affects the household's choice between consumption and leisure. If $\hat{\Lambda}_L$ is close to 1, the effect of wage changes on leisure is small, as most wage increases translate directly into price increases, reducing the real wage effect. The second term, the wealth effect, suggests that an increase in output would typically lead to an increase in leisure, reducing labor supply.

Proof of Proposition 2. By combining the previous equation with $\mathrm{d}\log Y - \mathrm{d}\log L = (1-\hat{\Lambda}_L - \xi)\mathrm{d}\log w$ from equation (5), we find that, following a change in nominal wage, the response of output is

$$\frac{\mathrm{d}\log Y}{\mathrm{d}\log w} = \frac{\mathrm{d}\log \mathrm{TFP}}{\mathrm{d}\log w} + \frac{\mathrm{d}\log L}{\mathrm{d}\log w} = \frac{1-\hat{\Lambda}_L + \varphi(1-\hat{\Lambda}_L-\xi)}{\gamma+\varphi}, \tag{D-8}$$

Specifically,

$$\frac{\mathrm{d}\log \mathrm{TFP}}{\mathrm{d}\log w} = 1-\hat{\Lambda}_L - \xi \quad \text{and} \quad \frac{\mathrm{d}\log L}{\mathrm{d}\log w} = \frac{1-\hat{\Lambda}_L - \gamma(1-\hat{\Lambda}_L-\xi)}{\gamma+\varphi}.$$

This analysis, in conjunction with equation (D-5), establishes Proposition 2. □

Equation (D-8) breaks down the output response into supply-side and demand-side effects. The supply-side effect is due to changes in TFP, arising through the misallocation channel. The demand-side effect characterizes the endogenous response of labor supply, which consists of two distinct channels: nominal rigidity and misallocation. Firstly, the substitution effect, highlighted in equation (D-7), illustrates the nominal rigidity channel, showing that an expansionary monetary shock enhances real wages and stimulates labor supply. Second, the wealth effect, also from equation (D-7), reveals that any factor that changes output, such as supply-side effects, also affects labor. Therefore, the demand-side effect in our model accounts for labor adjustments induced by the misallocation channel, quantified as $-\frac{\gamma}{\gamma+\varphi}(1-\hat{\Lambda}_L-\xi)$. Consequently, supply-side effects complement demand-side effects, leading to an increase in monetary non-neutrality.

E Optimal Monetary Policy

E.1 Discrepancy between Output and Employment Gaps

In this section, we aim to express aggregate macroeconomic variables in terms of changes in productivities and ex-post markups, and define output and employment gaps.

From the supply-side, we have

$$\pi = -\mathrm{d}\log A + \mathrm{d}\log\mu + \tilde{\Omega}^n\pi + \tilde{\Omega}^f\mathrm{d}\log w \tag{E-1}$$

$$= -\tilde{\Psi}^n\mathrm{d}\log A + \tilde{\Psi}^n\mathrm{d}\log\mu + \tilde{\Psi}^f\mathrm{d}\log w \tag{E-2}$$

$$= -\tilde{\Psi}^n\mathrm{d}\log A + \tilde{\Psi}^n\mathrm{d}\log\mu + \mathrm{d}\log w \tag{E-3}$$

Furthermore, the change in consumer price is

$$\mathrm{d}\log P^Y = b'\pi = -\tilde{\lambda}'\mathrm{d}\log A + \tilde{\lambda}'\mathrm{d}\log\mu + \mathrm{d}\log w \tag{E-4}$$

Combining with equation (C-29), this implies the change in TFP is determined by

$$\mathrm{d}\log Y - \mathrm{d}\log L = \tilde{\lambda}'\mathrm{d}\log A - \tilde{\lambda}'\mathrm{d}\log\mu - \mathrm{d}\log\Lambda_L. \tag{E-5}$$

This result aligns with proposition 2 of Baqaee and Farhi (2020).

On the other hand, from Theorem 1, the change in TFP relates with changes in productivities and changes in ex-post markups via

$$d\log Y - d\log L = \sum_{j=1}^{N} \lambda_j \frac{\Psi_{jL}}{\Lambda_L} d\log A_j + \sum_{j=0}^{N} \frac{\sigma_j \lambda_j}{\bar{\mu}_j} \mathrm{Cov}_{\tilde{\Omega}(j,:)}\left(d\log p_i, \frac{\Psi_{iL}}{\Lambda_L} \right) \tag{E-6}$$

$$= \sum_{j=1}^{N} \lambda_j \frac{\Psi_{jL}}{\Lambda_L} d\log A_j + \sum_{j=0}^{N} \frac{\sigma_j \lambda_j}{\bar{\mu}_j} \mathrm{Cov}_{\tilde{\Omega}(j,:)}\left(\sum_{k=1}^{N} \tilde{\Psi}_{ik}(d\log \mu_k - d\log A_k) + d\log w, \frac{\Psi_{iL}}{\Lambda_L} \right) \tag{E-7}$$

$$= \sum_{j=1}^{N} \lambda_j \frac{\Psi_{jL}}{\Lambda_L} d\log A_j + \sum_{k=1}^{N} \sum_{j=0}^{N} \frac{\sigma_j \lambda_j}{\bar{\mu}_j} \mathrm{Cov}_{\tilde{\Omega}(j,:)}\left(\tilde{\Psi}_{ik}, \frac{\Psi_{iL}}{\Lambda_L} \right) (d\log \mu_k - d\log A_k) \tag{E-8}$$

$$= \left[(\lambda \circ \frac{\Psi_{(L)}}{\Lambda_L})' - \mathcal{M}' \right] d\log A + \mathcal{M}' d\log \mu \tag{E-9}$$

where $\mathcal{M}$ is an $N \times 1$ vector, whose kth element is given by

$$\mathcal{M}_k \doteq \frac{d\log \mathrm{TFP}}{d\log \mu_k} = \sum_{j=0}^{N} \frac{\sigma_j \lambda_j}{\bar{\mu}_j} \mathrm{Cov}_{\tilde{\Omega}(j,:)}\left(\tilde{\Psi}_{ik}, \frac{\Psi_{iL}}{\Lambda_L} \right) \tag{E-10}$$

By combining equations (E-5) and (E-9), we obtain

$$d\log \Lambda_L = \left[\tilde{\lambda}' - (\lambda \circ \frac{\Psi_{(L)}}{\Lambda_L})' + \mathcal{M}' \right] d\log A - \left[\tilde{\lambda}' + \mathcal{M}' \right] d\log \mu \tag{E-11}$$

Combining the definition of labor income share (equation (C-29)) with the consumption-leisure trade-off (equation (D-1)), yields:

$$d\log L = \frac{1-\gamma}{1+\varphi} d\log Y + \frac{1}{1+\varphi} d\log \Lambda_L \tag{E-12}$$

Combined with equations (E-5) and (E-11), we get:

$$d\log Y = \frac{1+\varphi}{\gamma+\varphi} \tilde{\lambda}'(d\log A - d\log \mu) - \frac{\varphi}{\gamma+\varphi} d\log \Lambda_L \tag{E-13}$$

$$= \left\{ \frac{1}{\gamma+\varphi} \tilde{\lambda}' + \frac{\varphi}{\gamma+\varphi} \left[(\lambda \circ \frac{\Psi_{(L)}}{\Lambda_L})' - \mathcal{M}' \right] \right\} d\log A - \left[\frac{1}{\gamma+\varphi} \tilde{\lambda}' - \frac{\varphi}{\gamma+\varphi} \mathcal{M}' \right] d\log \mu \tag{E-14}$$

and

$$\mathrm{d}\log L = \frac{1-\gamma}{\gamma+\varphi}\tilde{\lambda}'(\mathrm{d}\log A - \mathrm{d}\log\mu) + \frac{\gamma}{\gamma+\varphi}\mathrm{d}\log\Lambda_L \tag{E-15}$$

$$= \left\{\frac{1}{\gamma+\varphi}\tilde{\lambda}' - \frac{\gamma}{\gamma+\varphi}\left[(\lambda\circ\frac{\Psi_{(L)}}{\Lambda_L})' - \mathcal{M}'\right]\right\}\mathrm{d}\log A - \left[\frac{1}{\gamma+\varphi}\tilde{\lambda}' + \frac{\gamma}{\gamma+\varphi}\mathcal{M}'\right]\mathrm{d}\log\mu. \tag{E-16}$$

In the flexible price equilibrium, changes in output and employment are given by

$$y^n \equiv \mathrm{d}\log Y^n = \left\{\frac{1}{\gamma+\varphi}\tilde{\lambda}' + \frac{\varphi}{\gamma+\varphi}\left[(\lambda\circ\frac{\Psi_{(L)}}{\Lambda_L})' - \mathcal{M}'\right]\right\}\mathrm{d}\log A \tag{E-17}$$

and

$$l^n \equiv \mathrm{d}\log L^n = \left\{\frac{1}{\gamma+\varphi}\tilde{\lambda}' - \frac{\gamma}{\gamma+\varphi}\left[(\lambda\circ\frac{\Psi_{(L)}}{\Lambda_L})' - \mathcal{M}'\right]\right\}\mathrm{d}\log A. \tag{E-18}$$

The output and employment gaps, which quantify the deviations between the sticky price and flexible price equilibria, can be expressed as:

$$\tilde{y} \equiv \mathrm{d}\log Y - \mathrm{d}\log Y^n = \left[-\frac{1}{\gamma+\varphi}\tilde{\lambda}' + \frac{\varphi}{\gamma+\varphi}\mathcal{M}'\right]\mathrm{d}\log\mu \tag{E-19}$$

and

$$\tilde{l} \equiv \mathrm{d}\log L - \mathrm{d}\log L^n = \left[-\frac{1}{\gamma+\varphi}\tilde{\lambda}' - \frac{\gamma}{\gamma+\varphi}\mathcal{M}'\right]\mathrm{d}\log\mu \tag{E-20}$$

When the initial equilibrium is inefficient, the discrepancy between output and employment gaps reflects an allocative efficiency:

$$e \equiv \tilde{y} - \tilde{l} = \mathcal{M}'\mathrm{d}\log\mu \tag{E-21}$$

Accounting for endogenous realized markups, the allocative efficiency e is related to sectoral inflation rates through the equation:

$$e = \underbrace{\mathcal{M}'(I-\Theta^{-1})}_{\mathcal{J}'}\pi \tag{E-22}$$

Specifically, in response to a change in nominal wages,

$$e = (1-\hat{\Lambda}_L - \xi)\mathrm{d}\log w = \mathcal{J}'\hat{\Psi}^f\mathrm{d}\log w. \tag{E-23}$$

E.2 Flatter Phillips Curves

Proof of Lemma 1. By combining equations (3) and (E-19), we derive the divine coincidence condition:

$$(\gamma+\varphi)\tilde{y} = \left[\tilde{\lambda}'(\Theta^{-1}-I)\underbrace{-\varphi\mathcal{M}'(\Theta^{-1}-I)}_{\varphi\mathcal{J}'}\right]\pi \tag{E-24}$$

$$= \left[\lambda'(\Theta^{-1}-I)+\varphi\mathcal{J}'\right]\pi \tag{E-25}$$

□

Combining with the sectoral inflation from equation (C-35), leads to a wage Phillips curve:

$$[1-\hat{\Lambda}_L+\varphi(1-\hat{\Lambda}_L-\xi)]\mathrm{d}\log w = (\gamma+\varphi)\tilde{y}+(\tilde{\lambda}'-\hat{\lambda}'\Theta+\varphi\mathcal{J}'\hat{\Psi}^n\Theta)\mathrm{d}\log A. \tag{E-26}$$

The derivation of this equation is based on three equations: (i) $\mathcal{J}'\hat{\Psi}^n_{(L)} = 1-\hat{\Lambda}_L-\xi$, (ii) $\tilde{\lambda}(\Theta-I)\hat{\Psi}^n_{(L)} = 1-\hat{\Lambda}_L$, and (iii) $\tilde{\lambda}(\Theta-I)\hat{\Psi}^n\Theta = \tilde{\lambda}'-\hat{\lambda}'\Theta$.

The wage Phillips curve also suggests that if the output response to a change in nominal wages is non-zero, $1-\hat{\Lambda}_L+\varphi(1-\hat{\Lambda}_L-\xi)\neq 0$, then a one-to-one mapping exists between nominal wage w and output gap $\tilde{y}$ for all realization of productivity shocks A. Consequently, combined with Lemma 6, this allows the monetary authority to target any desired output level by adjusting the money supply.

Proof of Proposition 3. Substituting the wage Phillips curve into sectoral inflation (equation (C-35)), yields sectoral Phillips curves,

$$\pi = -\hat{\Psi}^n\Theta\mathrm{d}\log A+\hat{\Psi}^f\mathrm{d}\log w \tag{E-27}$$

$$= \underbrace{\frac{\gamma+\varphi}{1-\hat{\Lambda}_L+\varphi(1-\hat{\Lambda}_L-\xi)}\hat{\Psi}^f}_{\mathcal{K}}\tilde{y} + \underbrace{\left[\frac{1}{1-\hat{\Lambda}_L+\varphi(1-\hat{\Lambda}_L-\xi)}\hat{\Psi}^f(\tilde{\lambda}'-\hat{\lambda}'\Theta+\varphi\mathcal{J}'\hat{\Psi}^n\Theta)-\hat{\Psi}^n\Theta\right]}_{\mathcal{V}}\mathrm{d}\log A \tag{E-28}$$

Note that $\left[\tilde{\lambda}'(\Theta^{-1} - I) + \varphi\mathcal{J}'\right]\mathcal{V} = \mathbf{0}$ and $\mathcal{V}\tilde{\Omega}^n_{(L)} = \mathbf{0}$. Given that $\tilde{\lambda}_i(\theta_i^{-1} - 1) + \varphi\mathcal{J}_i \geq 0$ and $\tilde{\Omega}_{iL} > 0$ for all $i \in \mathcal{N}$, these conditions guarantee that $\mathcal{V}$ is a non-zero matrix. [2] □

E.3 Optimal Monetary Policy in Distorted Economies

We now move to a second-order approximation around the flexible-price equilibrium with distortions. [3]

First, note that price dispersion within each sector is distortionary. Let q denote an $N \times 1$ vector representing the productivity loss due to within-sector price distortions, with component:

$$q_i \doteq \frac{y_i}{A_i F_i(\{x_{ij}\}_{j=1}^{N+1})} < 1 \tag{E-29}$$

where $x_{ij} = \int_0^1 x_{ij,\nu} \mathrm{d}\nu, \forall i, j$.

We then show that

$$q_i = \frac{p_i^{-\varepsilon_i}}{\int p_{i,\nu}^{-\varepsilon_i} d\nu}. \tag{E-30}$$

Cost minimization by the firm in sector i results in the following demand for input j:

$$x_{ij,\nu} = A_i^{\sigma_i - 1}\omega_{ij} y_{i,\nu}\left(p_j/mc_i\right)^{-\sigma_i}, \quad \forall i, j \tag{E-31}$$

Hence, the aggregate demand for input j by firms in industry i is

$$x_{ij} = \int_0^1 x_{ij,\nu} \mathrm{d}\nu \tag{E-32}$$

$$= A_i^{\sigma_i - 1}\omega_{ij} \int_0^1 y_{i,\nu} \mathrm{d}\nu \left(p_j/mc_i\right)^{-\sigma_i} \tag{E-33}$$

$$= A_i^{\sigma_i - 1}\omega_{ij} y_i \left(p_j/mc_i\right)^{-\sigma_i} p_i^{\varepsilon_i} \int_0^1 p_{i,\nu}^{-\varepsilon_i} \mathrm{d}\nu \tag{E-34}$$

where the last equality uses the fact that $y_{i,\nu} = y_i(p_{i,\nu}/p_i)^{-\varepsilon_i}$.

---

[2] Within the divine coincidence inflation index, in general the term $\tilde{\lambda}_i(\theta_i^{-1} - 1)$ is dominant and positive.

[3] A second-order approximation of a variable Z around its deterministic steady state Z^* is written as,

$$\frac{Z - Z^*}{Z^*} \approx \hat{z} + \frac{1}{2}\hat{z}^2$$

where $\hat{z} = \Delta \log z = \log z - \log Z^*$.

Then, we have

$$q_i = \frac{y_i}{A_i\left(\sum_j \omega_{ij}^{\frac{1}{\sigma_i}} x_{ij}^{\frac{\sigma_i-1}{\sigma_i}}\right)^{\frac{\sigma_i}{\sigma_i-1}}} = \frac{mc_i^{-\sigma_i}}{\underbrace{\left[\frac{1}{A_i}\left(\sum_j \omega_{ij} p_j^{1-\sigma_i}\right)^{\frac{1}{1-\sigma_i}}\right]^{-\sigma_i}}_{=1}} \frac{p_i^{-\varepsilon_i}}{\int p_{i,v}^{-\varepsilon_i} dv} \tag{E-35}$$

In line with the traditional NK model (Galí, 2015), we observe $\mathrm{d}\log q_i = 0$ and:

$$-d^2 \log q_i = \varepsilon_i \mathrm{Var}_i(p_{i,v}) \tag{E-36}$$

$$= \varepsilon_i \left[\int (\log p_{i,v} - \log p_i)^2 dv - \left(\int (\log p_{i,v} - \log p_i) dv\right)^2\right] \tag{E-37}$$

$$= \varepsilon_i \left(\frac{1}{\theta_i} - 1\right)(\mathrm{d}\log p_i)^2. \tag{E-38}$$

Lemma 7. *Up to a second-order approximation, the logarithmic change in output per labor in the sticky price equilibrium is given by*

$$\hat{y} - \hat{l} = \underbrace{e}_{\textit{first-order}} - \underbrace{f}_{\textit{second order}} + \textit{higher order terms} \tag{E-39}$$

where e is allocative efficiency given by equation (E-21).

Proof of Lemma 7. Following the same steps as in the proof of Theorem 1, we can derive

$$\mathrm{d}\log Y - \mathrm{d}\log L = \left[(\lambda \circ \frac{\Psi_{(L)}}{\Lambda_L})' - \mathcal{M}'\right](\mathrm{d}\log A + \mathrm{d}\log q) + \mathcal{M}'\mathrm{d}\log \mu \tag{E-40}$$

and

$$\mathrm{d}\log Y^n - \mathrm{d}\log L^n = \left[(\lambda \circ \frac{\Psi_{(L)}}{\Lambda_L})' - \mathcal{M}'\right]\mathrm{d}\log A \tag{E-41}$$

A first-order approximation of the logarithmic change in output per labor is

$$\mathrm{d}(\hat{y} - \hat{l}) = (\mathrm{d}\log Y - \mathrm{d}\log Y^n) - (\mathrm{d}\log L - \mathrm{d}\log L^n) \tag{E-42}$$

$$= (\mathrm{d}\log Y - \mathrm{d}\log L) - (\mathrm{d}\log Y^n - \mathrm{d}\log L^n) \tag{E-43}$$

$$= \left[(\lambda \circ \frac{\Psi_{(L)}}{\Lambda_L})' - \mathcal{M}'\right]\mathrm{d}\log q + \mathcal{M}'\mathrm{d}\log \mu \tag{E-44}$$

$$= \mathcal{M}'\mathrm{d}\log \mu = e \tag{E-45}$$

Differentiating the equation (E-44) again, we obtain

$$\mathrm{d}^2(\hat{y}-\hat{l}) = \left[(\lambda \circ \frac{\Psi_{(L)}}{\Lambda_L})' - \mathcal{M}'\right]\mathrm{d}^2 \log q + \sum_i \sum_j \frac{\mathrm{d}\log \mathcal{M}_i}{\mathrm{d}\log \mu_j}\mathrm{d}\log \mu_i \mathrm{d}\log \mu_j \tag{E-46}$$

Thus, the second-order component f, consists of two critical components: within-sector misallocation ($\mathcal{L}^{\text{within}}$) and cross-sector misallocation ($\mathcal{L}^{\text{across}}$)

$$f = \mathcal{L}^{\text{within}} + \mathcal{L}^{\text{across}}. \tag{E-47}$$

The second-order welfare loss due to within-sector misallocation can be expressed as

$$\mathcal{L}^{\text{within}} = -\frac{1}{2}\sum_i \left(\lambda_i \frac{\Psi_{iL}}{\Lambda_L} - \mathcal{M}_i\right)\mathrm{d}^2 \log q_i \doteq \frac{1}{2}\pi' \mathcal{H}_1 \pi, \tag{E-48}$$

where $\mathcal{H}_1 = \mathrm{diag}((\lambda \circ \frac{\Psi_{iL}}{\Lambda_L} - \mathcal{M}) \circ \epsilon \circ (\theta^{-1}-1))$.

The cross-sector misallocation is determined by:

$$\mathcal{L}^{\text{across}} = -\frac{1}{2}\sum_i \sum_j \frac{\mathrm{d}\log \mathcal{M}_i}{\mathrm{d}\log \mu_j}\mathrm{d}\log \mu_i \mathrm{d}\log \mu_j \doteq \frac{1}{2}\pi' \mathcal{H}_2 \pi. \tag{E-49}$$

Referring to Baqaee and Farhi (2020), a second-order approximation of the cross-sector misallocation is given by:

$$\mathcal{L}^{\text{across}} \approx \frac{1}{2}\sum_j \sigma_j \lambda_j \mathrm{Var}_{\Omega(j,:)}(\sum_k \Psi_{(k)} \mathrm{d}\log \mu_k) \tag{E-50}$$

which implies that

$$\mathcal{L}^{\text{across}} \approx \underbrace{\frac{1}{2}\sum_j \sigma_j \lambda_j \mathrm{Cov}_{\Omega(j,:)}(\Psi_{(k)}, \Psi_{(l)})}_{\doteq \mathcal{B}(k,l)} \mathrm{d}\log \mu_k \mathrm{d}\log \mu_l = \frac{1}{2}\pi(I-\Theta^{-1})\mathcal{B}(I-\Theta^{-1})\pi \tag{E-51}$$

Hence, $\mathcal{H}_2$ is given by $(I-\Theta^{-1})\mathcal{B}(I-\Theta^{-1})$.

□

Proof of Proposition 4. Using Lemma 7, we can approximate the utility function around the flexible price equilibrium as [4]

$$\frac{U-U^n}{U_y Y} \approx \hat{y} + \frac{1}{2}\hat{y}^2 + \frac{1}{2}\frac{U_{yy}Y}{U_y}\hat{y}^2 + \frac{U_l L}{U_y Y}(\hat{l} + \frac{1}{2}\hat{l}^2 + \frac{1}{2}\frac{U_{ll}L}{U_l}\hat{l}^2) \tag{E-52}$$

$$= \hat{y} + \frac{1-\gamma}{2}\hat{y}^2 - \Lambda_L(\hat{l} + \frac{1+\varphi}{2}\hat{l}^2) \tag{E-53}$$

$$= \hat{y} + \frac{1-\gamma}{2}\hat{y}^2 - \Lambda_L\left(\hat{y} - e + f + \frac{1+\varphi}{2}(\hat{y} - e + f)^2\right) \tag{E-54}$$

$$= (1-\Lambda_L)\hat{y} + \Lambda_L(e-f) + \frac{1-\gamma}{2}\hat{y}^2 - \frac{1+\varphi}{2}\Lambda_L(\hat{y}-e)^2 - \underbrace{\frac{1+\varphi}{2}\Lambda_L f(2\hat{y} - 2e + f)}_{O((\mathrm{d}\log\mu)^3)} \tag{E-55}$$

$$\approx (1-\Lambda_L)\hat{y} + \Lambda_L(e-f) + \frac{1-\gamma}{2}\hat{y}^2 - \frac{1+\varphi}{2}\Lambda_L(\hat{y}-e)^2 \tag{E-56}$$

$$= (1-\Lambda_L)\hat{y} + \Lambda_L e - f + \frac{1-\gamma}{2}\hat{y}^2 - \frac{1+\varphi}{2}(\hat{y}-e)^2 + \underbrace{(1-\Lambda_L)[f + \frac{1+\varphi}{2}(\hat{y}-e)^2]}_{O((\mathrm{d}\log\mu)^3)} \tag{E-57}$$

$$\approx (1-\Lambda_L)\hat{y} + \Lambda_L e - f + \frac{1-\gamma}{2}\hat{y}^2 - \frac{1+\varphi}{2}(\hat{y}-e)^2 \tag{E-58}$$

$$= (1-\Lambda_L)\hat{y} + \Lambda_L e - f - \frac{\gamma+\varphi}{2}(\hat{y} - \frac{1+\varphi}{\gamma+\varphi}e)^2 - \frac{\gamma-1}{2}\frac{1+\varphi}{\gamma+\varphi}e^2 \tag{E-59}$$

$$\approx (1-\Lambda_L)\tilde{y} + \Lambda_L e - f - \frac{\gamma+\varphi}{2}(\tilde{y} - \frac{1+\varphi}{\gamma+\varphi}e)^2 - \frac{\gamma-1}{2}\frac{1+\varphi}{\gamma+\varphi}e^2$$

$$+ \underbrace{(\hat{y}-\tilde{y})}_{O((\mathrm{d}\log\mu)^2)} \underbrace{\left[(1-\Lambda_L) - \frac{\gamma+\varphi}{2}(\hat{y} + \tilde{y} - 2\frac{1+\varphi}{\gamma+\varphi}e)\right]}_{O(\mathrm{d}\log\mu)} \tag{E-60}$$

$$\approx (1-\Lambda_L)\tilde{y} + \Lambda_L e - f - \frac{\gamma+\varphi}{2}(\tilde{y} - \frac{1+\varphi}{\gamma+\varphi}e)^2 - \frac{\gamma-1}{2}\frac{1+\varphi}{\gamma+\varphi}e^2. \tag{E-61}$$

In this derivation, we uses the fact that the difference between $\hat{y}$ and $\tilde{y}$ is given by second order terms:

$$\hat{y} - \tilde{y} \approx \frac{1}{2}\left\{\frac{1}{\gamma+\varphi}\tilde{\lambda}' + \frac{\varphi}{\gamma+\varphi}\left[(\lambda \circ \frac{\Psi_{(L)}}{\Lambda_L})' - \mathcal{M}'\right]\right\}\mathrm{d}^2\log q$$
$$+ \frac{1}{2}\frac{\varphi}{\gamma+\varphi}\sum_i\sum_j \frac{\mathrm{d}\log\mathcal{M}_i}{\mathrm{d}\log\mu_j}\mathrm{d}\log\mu_i \mathrm{d}\log\mu_j \tag{E-62}$$

Substituting $e = \mathcal{J}'\pi$ and $f = \frac{1}{2}\pi'\mathcal{H}_1\pi + \frac{1}{2}\pi'\mathcal{H}_2\pi$ into equation (E-61), the welfare function

[4] Under the assumption of small distortions in the equilibrium, the product of $1-\Lambda_L$ with a second-order term is a third-order timer and can be dropped from the approximation.

is then given by

$$\mathbb{W} = \underbrace{(1-\Lambda_L)\tilde{y} + \Lambda_L \mathcal{J}'\pi}_{\text{first-order bias}} \quad \underbrace{-\frac{\gamma+\varphi}{2}(\tilde{y} - \frac{1+\varphi}{\gamma+\varphi}\mathcal{J}'\pi)^2}_{\text{Volatility of output gap}} \quad \underbrace{-\frac{1}{2}\pi'\mathcal{H}_1\pi}_{\text{Within-sector price dispersion}}$$
$$\underbrace{-\frac{1}{2}\pi'\mathcal{H}_2\pi}_{\text{Cross-sector price dispersion}} \quad \underbrace{-\frac{\gamma-1}{2}\frac{1+\varphi}{\gamma+\varphi}\pi'\mathcal{J}\mathcal{J}'\pi}_{\text{Variation in allocative efficiency}}. \tag{E-63}$$

□

Proof of Proposition 5. The optimal monetary policy problem can be written as

$$\max_{\tilde{y},\pi} \mathbb{W} = (1-\Lambda_L)\tilde{y} + \Lambda_L\mathcal{J}'\pi - \frac{\gamma+\varphi}{2}\left(\tilde{y} - \frac{1+\varphi}{\gamma+\varphi}\mathcal{J}'\pi\right)^2 - \frac{1}{2}\pi'\mathcal{H}\pi - \frac{\gamma-1}{2}\frac{1+\varphi}{\gamma+\varphi}\pi'\mathcal{J}\mathcal{J}'\pi$$

subject to

$$\pi = \mathcal{K}\tilde{y} + \mathcal{V}\,\mathrm{d}\log A.$$

The Lagrangian $\mathcal{L}$ is

$$\mathcal{L}(\tilde{y},\pi;\psi) = \mathbb{W}(\tilde{y},\pi) - \psi'(\pi - \mathcal{K}\tilde{y} - \mathcal{V}\,\mathrm{d}\log A) \tag{E-64}$$

The corresponding first-order conditions are

$$(1-\Lambda_L) - (\gamma+\varphi)\left(\tilde{y} - \frac{1+\varphi}{\gamma+\varphi}\mathcal{J}'\pi\right) + \psi'\mathcal{K} = 0 \tag{E-65}$$

and

$$\Lambda_L\mathcal{J}' + (1+\varphi)\left(\tilde{y} - \frac{1+\varphi}{\gamma+\varphi}\mathcal{J}'\pi\right)\mathcal{J}' - \pi'\mathcal{H} - (\gamma-1)\frac{1+\varphi}{\gamma+\varphi}\pi'\mathcal{J}\mathcal{J}' - \psi' = 0. \tag{E-66}$$

Combine two equations above and obtain

$$\Lambda_L\mathcal{J}' + \frac{1+\varphi}{\gamma+\varphi}(1-\Lambda_L+\psi'\mathcal{K})\mathcal{J}' - \pi'\mathcal{H} - (\gamma-1)\frac{1+\varphi}{\gamma+\varphi}\pi'\mathcal{J}\mathcal{J}' - \psi' = 0, \tag{E-67}$$

or equivalently,

$$\psi' = \left[\Lambda_L \mathcal{J}' + \frac{1+\varphi}{\gamma+\varphi}(1-\Lambda_L)\mathcal{J}' - \pi'\mathcal{H} - (\gamma-1)\frac{1+\varphi}{\gamma+\varphi}\pi'\mathcal{J}\mathcal{J}'\right]\left[I - \frac{1+\varphi}{\gamma+\varphi}\mathcal{K}\mathcal{J}'\right]^{-1}. \tag{E-68}$$

To simplify the equation, we observe that

$$\left[I - \frac{1+\varphi}{\gamma+\varphi}\mathcal{K}\mathcal{J}'\right]^{-1} = I + \frac{(1+\varphi)\hat{\Psi}^f \mathcal{J}'}{(1-\hat{\Lambda}_L) + \varphi(1-\hat{\Lambda}_L-\xi) - (1+\varphi)\mathcal{J}'\hat{\Psi}^f} = I + \frac{1+\varphi}{\xi}\hat{\Psi}^f\mathcal{J}' \tag{E-69}$$

and

$$\left[I - \frac{1+\varphi}{\gamma+\varphi}\mathcal{K}\mathcal{J}'\right]^{-1}\mathcal{K} = \frac{\gamma+\varphi}{(1-\hat{\Lambda}_L)+\varphi(1-\hat{\Lambda}_L-\xi)}\left[\hat{\Psi}^f + \frac{1+\varphi}{\xi}\hat{\Psi}^f\mathcal{J}'\hat{\Psi}^f\right] = \frac{\gamma+\varphi}{\xi}\hat{\Psi}^f. \tag{E-70}$$

Hence, combining equations (E-65) and (E-68) yields

$$(\gamma+\varphi)(\tilde{y} - \frac{1+\varphi}{\gamma+\varphi}\mathcal{J}'\pi) + \frac{\gamma+\varphi}{\xi}(\hat{\Psi}^f)'\mathcal{H}\pi + \frac{(1+\varphi)(\gamma-1)(1-\hat{\Lambda}_L-\xi)}{\xi}\mathcal{J}'\pi \tag{E-71}$$

$$=1-\Lambda_L + [1+\varphi+(\gamma-1)\Lambda_L](1-\hat{\Lambda}_L-\xi)/\xi \tag{E-72}$$

Using divine coincidence condition, we get

$$\left[\tilde{\lambda}'(\Theta^{-1}-I) - \mathcal{J}' + \frac{\gamma+\varphi}{\xi}(\hat{\Psi}^f)'\mathcal{H} + \frac{(1+\varphi)(\gamma-1)(1-\hat{\Lambda}_L-\xi)}{\xi}\mathcal{J}'\right]\pi$$
$$=1-\Lambda_L + [1+\varphi+(\gamma-1)\Lambda_L](1-\hat{\Lambda}_L-\xi)/\xi \tag{E-73}$$

Note that when the flexible price equilibrium is efficient, $\Lambda_L = 1$, $\mathcal{J} = \mathbf{0}$ and $1-\hat{\Lambda}_L-\xi = 0$. The condition above degenerates to the results of La'O and Tahbaz-Salehi (2022) and Rubbo (2023):

$$\left[\tilde{\lambda}'(\Theta^{-1}-I) + \frac{\gamma+\varphi}{1-\hat{\Lambda}_L}(\hat{\Psi}^f)'\mathcal{H}\right]\pi = 0$$

□

F Cobb-Douglas Economy

In this section, we explore a Cobb-Douglas model economy characterized by unitary elasticities of substitution in both consumption and production ($\sigma_i = 1, \forall i$).

F.1 Supply-Side Effect in a Cobb-Douglass Economy

With unitary elasticity of substitution, the supply-side effect of monetary policy in Proposition 1 can be simplified as,

$$\frac{\mathrm{d}\log \mathrm{TFP}}{\mathrm{d}\log w} = \sum_{j=0}^{N} \frac{\lambda_j}{\bar{\mu}_j} \mathrm{Cov}_{\tilde{\Omega}(j,:)}\left(\hat{\Psi}_{iL}, \frac{\Psi_{iL}}{\Lambda_L}\right) \tag{F-1}$$

$$= 1 - \hat{\Lambda}_L - \sum_{j=1}^{N} \lambda_j \frac{\Psi_{iL}}{\Lambda_L} \frac{1-\theta_j}{\theta_j} \hat{\Psi}_{jL} \tag{F-2}$$

$$= \sum_{j=1}^{N} \tilde{\lambda}_j (1 - \frac{\lambda_j}{\tilde{\lambda}_j} \frac{\Psi_{iL}}{\Lambda_L}) \frac{1-\theta_j}{\theta_j} \hat{\Psi}_{jL} \tag{F-3}$$

where the second equality is obtained by taking directive of both sides of equation (C-27) with respect to $\mathrm{d}\log w$, and the last equality uses the fact that

$$1 - \hat{\Lambda}_L = \sum_{j=1}^{N} \tilde{\lambda}_j \frac{1-\theta_j}{\theta_j} \hat{\Psi}_{jL}. \tag{F-4}$$

Alternatively, the sufficient statistic for the supply-side effect can be derived through the following lemma.

Lemma 8. *In a Cobb-Douglas economy, the change in TFP, in terms of productivities and ex-post markups, is determined by the following expression:*

$$\mathrm{d}\log \mathrm{TFP} = \tilde{\lambda}' \mathrm{d}\log A + \left[(\lambda \circ \frac{\Psi_{(L)}}{\Lambda_L})' - \tilde{\lambda}' \right] \mathrm{d}\log \mu \tag{F-5}$$

Proof of Lemma 8. When all cross-sector elasticities are set to one, we have

$$\mathcal{M}_k = \sum_{j=0}^{N} \frac{\lambda_j}{\bar{\mu}_j} \mathrm{Cov}_{\tilde{\Omega}(j,:)}\left(\tilde{\Psi}_{ik}, \frac{\Psi_{iL}}{\Lambda_L}\right) \tag{F-6}$$

$$= \sum_{j=0}^{N} \frac{\lambda_j}{\bar{\mu}_j} \left[\sum_i \tilde{\Omega}_{ji} \tilde{\Psi}_{ik} \frac{\Psi_{iL}}{\Lambda_L} - \underbrace{\left(\sum_i \tilde{\Omega}_{ji} \tilde{\Psi}_{ik} \right)}_{\tilde{\Psi}_{jk} - \delta_{jk}} \underbrace{\left(\sum_i \tilde{\Omega}_{ji} \frac{\Psi_{iL}}{\Lambda_L}) \right)}_{\bar{\mu}_j \Psi_{jL}/\Lambda_L} \right] \tag{F-7}$$

$$= \sum_{j=0}^{N} \lambda_j \sum_{i=1}^{N+1} \Omega_{ji} \tilde{\Psi}_{ik} \frac{\Psi_{iL}}{\Lambda_L} - \sum_{j=0}^{N} \lambda_j (\tilde{\Psi}_{jk} - \delta_{jk}) \frac{\Psi_{jL}}{\Lambda_L} \tag{F-8}$$

$$= \sum_{i=1}^{N+1} \tilde{\Psi}_{ik} \frac{\Psi_{iL}}{\Lambda_L} \underbrace{\sum_{j=0}^{N} \lambda_j \Omega_{ji}}_{\lambda_i} - \sum_{j=0}^{N} \lambda_j \tilde{\Psi}_{jk} \frac{\Psi_{jL}}{\Lambda_L} + \lambda_k \frac{\Psi_{kL}}{\Lambda_L} \tag{F-9}$$

$$= \sum_{i=1}^{N} \lambda_i \tilde{\Psi}_{ik} \frac{\Psi_{iL}}{\Lambda_L} - \sum_{j=0}^{N} \lambda_j \tilde{\Psi}_{jk} \frac{\Psi_{jL}}{\Lambda_L} + \lambda_k \frac{\Psi_{kL}}{\Lambda_L} \tag{F-10}$$

$$= - \underbrace{\lambda_0 \tilde{\Psi}_{0k} \frac{\Psi_{0L}}{\Lambda_L}}_{\tilde{\lambda}_k} + \lambda_k \frac{\Psi_{kL}}{\Lambda_L} \tag{F-11}$$

$$= \lambda_k \frac{\Psi_{kL}}{\Lambda_L} - \tilde{\lambda}_k \tag{F-12}$$

This, combining with equation (E-9), completes the proof. □

By Lemma 8, we obtain

$$\frac{\mathrm{d} \log \mathrm{TFP}}{\mathrm{d} \log w} = \left[(\lambda \circ \frac{\Psi_{(L)}}{\Lambda_L})' - \tilde{\lambda}' \right] (I - \Theta^{-1}) \hat{\Psi}^f \tag{F-13}$$

$$= \sum_{j=1}^{N} \tilde{\lambda}_j (1 - \frac{\lambda_j}{\tilde{\lambda}_j} \frac{\Psi_{iL}}{\Lambda_L}) \frac{1 - \theta_j}{\theta_j} \hat{\Psi}_{jL} \tag{F-14}$$

F.2 Network-Adjusted Markups

Inspired by equation (F-14), we define χ_i as a network-adjusted markup. Specifically, for all $i \in \mathcal{N}$, it is given by:

$$\chi_i = \left(\frac{\lambda_i}{\tilde{\lambda}_i}\frac{\Psi_{iL}}{\Lambda_L}\right)^{-1}, \tag{F-15}$$

where $\frac{\lambda_i}{\tilde{\lambda}_i}$ captures sector i's downstream markup, while $\frac{\Psi_{iL}}{\Lambda_L}$ represents sector i's upstream markup relative to the aggregate wedge. Therefore, their product $\frac{\lambda_i}{\tilde{\lambda}_i}\frac{\Psi_{iL}}{\Lambda_L}$ captures the network-adjusted markup of sector i. [5]

With this definition, equation (F-14) simplifies to

$$\frac{\mathrm{d}\log \mathrm{TFP}}{\mathrm{d}\log w} = \sum_{j=1}^{N}(1-\chi_j^{-1})\tilde{\lambda}_j\frac{1-\theta_j}{\theta_j}\hat{\Psi}_{jL}. \tag{F-16}$$

This equation illustrates that, in a Cobb-Douglas economy, the supply-side effect depends on the interaction between network-adjusted markups measured by $1-\chi_i^{-1}$ and network-adjusted nominal rigidities represented by $(\theta_i^{-1}-1)\hat{\Psi}_{iL}$. Without either initial distortions ($\chi_i = 1, \forall i \in \mathcal{N}$), or nominal rigidities ($\theta = \mathbf{1}$), there is no supply-side effect.

Note that if a sector has a higher markup, and its upstream and downstream sectors also have higher markups, it is generally associated with higher network-adjusted markups (higher χ_i). Specifically, when two sectors i and j are symmetric both upstream and downstream, that is, they share identical production technologies and have the same roles as input suppliers to firms and in household preferences (La'O and Tahbaz-Salehi, 2022), their network-adjusted markups, χ_i and χ_j, directly reflect their respective markups, $\bar{\mu}_i$ and $\bar{\mu}_j$, as demonstrated in Lemma 9.

Lemma 9. *If sector i and j are upstream and downstream symmetric, then $\chi_i > \chi_j$ if and only if $\bar{\mu}_i > \bar{\mu}_j$.*

Proof of Lemma 9. Given sector i and j are upstream and downstream symmetric, we have $\tilde{\lambda}_i = \tilde{\lambda}_j$. Hence, $\left(\frac{\lambda_i}{\tilde{\lambda}_i}\frac{\Psi_{iL}}{\Lambda_L}\right)^{-1} > \left(\frac{\lambda_j}{\tilde{\lambda}_j}\frac{\Psi_{jL}}{\Lambda_L}\right)^{-1}$ is equivalent to $\lambda_i\Psi_{iL} < \lambda_j\Psi_{jL}$.

[5] For example, in a vertical economy, the downstream and upstream markups are given by,

$$\frac{\lambda_i}{\tilde{\lambda}_i} = \prod_{k=1}^{i-1}\bar{\mu}_k^{-1} \quad \text{and} \quad \frac{\Psi_{iL}}{\Lambda_L} = \frac{\prod_{k=i}^{N}\bar{\mu}_k^{-1}}{\prod_{k=1}^{N}\bar{\mu}_k^{-1}} = \prod_{k=1}^{i-1}\bar{\mu}_k.$$

Here, downstream and upstream markups exactly cancel each other out, resulting in network-adjusted markups uniformly being equal to one.

Similar to Lemma 4, we find:

$$\frac{\partial \log \lambda_k}{\partial \log \bar{\mu}_k} = -(\Psi_{kk} - 1) \leq 0 \tag{F-17}$$

and

$$\frac{\partial \log \Psi_{kL}}{\partial \log \bar{\mu}_k} = -\Psi_{kk} < 0. \tag{F-18}$$

Therefore, $\lambda_i \Psi_{iL} < \lambda_j \Psi_{jL}$ is also equivalent to $\bar{\mu}_i > \bar{\mu}_j$.

□

F.3 Optimal Monetary Policy in a Cobb-Douglas Economy

Following La'O and Tahbaz-Salehi (2022), we rewrite the optimal policy weight for industry i, ϕ_i^*, as a sum of components related to the output gap ($o.g.$), within-sector price dispersion ($within$), cross-sector price dispersion ($across$), and variations in allocative efficiency ($adjust$):

$$\phi_i^* \equiv \phi_i^{o.g.} + \phi_i^{within} + \phi_i^{across} + \phi_i^{adjust}. \tag{F-19}$$

To investigate the optimal monetary policy in a Cobb-Douglas economy, we begin by analyzing matrices $\mathcal{H}_1$ and $\mathcal{H}_2$, which are associated with within-sector and cross-sector misallocations, respectively.

Within-Sector Misallocation

The matrix $\mathcal{H}_1$ associated with within-sector misallocation is

$$\mathcal{H}_1 = \text{diag}(\tilde{\lambda} \circ \epsilon \circ (\theta^{-1} - 1)) \tag{F-20}$$

The derivation of this equation is based on Lemma 8, which states $\mathcal{M}_k = \lambda_k \frac{\Psi_{kL}}{\Lambda_L} - \tilde{\lambda}_k$.

Cross-Sector Misallocation

In terms of cross-sector misallocations, we deduce from Lemma 10 that for Cobb-Douglas elasticities across all sectors.

Lemma 10. *When elasticities of substitution are Cobb-Douglas across all sectors, $\sigma_j = 1, \forall j$, then*

$$\mathcal{B}(k,l) = \sum_j \lambda_j \text{Cov}_{\Omega(j,:)}(\Psi_{(k)}, \Psi_{(l)}) = \lambda_k \lambda_l \left[\frac{\Psi_{lk}}{\lambda_k} + \frac{\Psi_{kl}}{\lambda_l} - \frac{\delta_{kl}}{\lambda_k} - 1 \right] \tag{F-21}$$

Proof of Lemma 10.

$$\sum_j \lambda_j \mathrm{Cov}_{\Omega(j,:)}(\Psi_{ik}, \Psi_{il}) = \sum_j \lambda_j \left[\sum_i \Omega_{ji}\Psi_{ik}\Psi_{il} - \underbrace{\left(\sum_i \Omega_{ji}\Psi_{ik}\right)}_{\Psi_{jk}-\delta_{jk}} \underbrace{\left(\sum_i \Omega_{ji}\Psi_{il}\right)}_{\Psi_{jl}-\delta_{jl}} \right] \tag{F-22}$$

$$= \sum_{i=1}^{N+1} \Psi_{ik}\Psi_{il} \underbrace{\sum_{j=0}^{N} \lambda_j \Omega_{ji}}_{\lambda_i} - \sum_{j=0}^{N} \lambda_j (\Psi_{jk} - \delta_{jk})(\Psi_{jl} - \delta_{jl}) \tag{F-23}$$

$$= \sum_{i=1}^{N} \lambda_i \Psi_{ik}\Psi_{il} - \sum_{j=1}^{N} \lambda_j (\Psi_{jk} - \delta_{jk})(\Psi_{jl} - \delta_{jl}) - \underbrace{\lambda_0 \Psi_{0k}\Psi_{0l}}_{\lambda_k \lambda_l} \tag{F-24}$$

$$= \lambda_k \lambda_l \left[\frac{\Psi_{lk}}{\lambda_k} + \frac{\Psi_{kl}}{\lambda_l} - \frac{\delta_{kl}}{\lambda_k} - 1 \right] \tag{F-25}$$

□

In this paper, we refine our understanding of cross-sector misallocation in a Cobb-Douglas economy. We derive the following expression for $\mathcal{L}^{\text{across}}$:

$$\mathcal{L}^{\text{across}} = \sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_j \Psi_{ji} \frac{\Psi_{iL}}{\Lambda_L} \mathrm{d}\log\mu_i \mathrm{d}\log\mu_j - \frac{1}{2}\sum_{i=1}^{N} \lambda_i \frac{\Psi_{iL}}{\Lambda_L} \mathrm{d}\log^2\mu_i - \frac{1}{2}\left(\sum_{i=1}^{N} \lambda_i \frac{\Psi_{iL}}{\Lambda_L} \mathrm{d}\log\mu_i\right)^2 \tag{F-26}$$

$$= \sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_k \lambda_l \left[\frac{\Psi_{lk}}{\lambda_k}\frac{\Psi_{kL}}{\Lambda_L} + \frac{\Psi_{kl}}{\lambda_l}\frac{\Psi_{lL}}{\Lambda_L} - \frac{\delta_{kl}}{\lambda_k}\frac{\Psi_{kL}}{\Lambda_L} - \frac{\Psi_{kL}}{\Lambda_L}\frac{\Psi_{lL}}{\Lambda_L} \right] \mathrm{d}\log\mu_k \mathrm{d}\log\mu_l \tag{F-27}$$

This leads to the expression for the matrix $\mathcal{H}_2$ associated with cross-sector misallocation:

$$\mathcal{H}_2 = (I - \Theta^{-1})\mathcal{B}(I - \Theta^{-1}), \tag{F-28}$$

where $\mathcal{B} = \mathcal{D} + \mathcal{D}' - \mathrm{diag}(\lambda \circ \frac{\Psi^f}{\Lambda_L}) - (\lambda \circ \frac{\Psi^f}{\Lambda_L})(\lambda \circ \frac{\Psi^f}{\Lambda_L})'$ and $\mathcal{D} = \mathrm{diag}(\lambda)\Psi^n \mathrm{diag}(\frac{\Psi^f}{\Lambda_L})$.

Finally, we derive the components of the inflation index under optimal monetary policy, which can be simplified as follows:

$$\phi_i^{o.g.} = (\theta_i^{-1} - 1)\lambda_i \Psi_{iL}/\Lambda_L = (\theta_i^{-1} - 1)\tilde{\lambda}_i \chi_i^{-1}, \tag{F-29}$$

$$\phi_i^{within} = \frac{\gamma + \varphi}{\xi}(\theta_i^{-1} - 1)\tilde{\lambda}_i \varepsilon_i \hat{\Psi}_{iL}, \tag{F-30}$$

$$\phi_i^{across} = \frac{\gamma + \varphi}{\xi}(\theta_i^{-1} - 1)\left\{\sum_{k=1}^{N}(\theta_k^{-1} - 1)\hat{\Psi}_{kL}(\lambda_k \Psi_{ki}\Psi_{iL} + \lambda_i \Psi_{ik}\Psi_{kL}) - \lambda_i \Psi_{iL}[\hat{\Psi}_{iL}(\theta_i^{-1} - 1) + \xi]\right\}/\Lambda_L, \tag{F-31}$$

$$\phi_i^{adjust} = \frac{(1 + \varphi)(\gamma - 1)(1 - \hat{\Lambda}_L - \xi)}{\xi}(\theta_i^{-1} - 1)\tilde{\lambda}_i(1 - \chi_i^{-1}), \tag{F-32}$$

In a Cobb-Douglas economy, the output-gap stabilization policy is equivalent to $\sum_i \phi_i^{o.g.}\pi = 0$, with $\phi^{o.g.} = (\theta_i^{-1} - 1)\lambda_i \chi_i^{-1}$, indicating that the monetary authority could utilize allocative efficiency by allowing for higher inflation in sectors with higher network-adjusted markups (higher χ_i). Moreover, Lemma 9 indicates that sectors with higher network-adjusted markups generally correspond to lower initial markups.

G Optimal Policy in Example Economies

Example 1. Vertical Economy

The optimal monetary policy in a vertical economy can be simplified as

$$\left[\underbrace{\frac{\gamma + \varphi}{1 - \hat{\Lambda}_L}(\hat{\Psi}^n_{(L)})'\mathcal{H}_1}_{\text{Within-sector price dispersion}} + \underbrace{\Theta^{-1} - I}_{\text{Output gap}}\right]\pi = 1 - \prod_i \bar{\mu}_i^{-1}$$

In a vertical economy, cross-sectional resource allocation is always efficient regardless of initial wedges and pricing rigidities, therefore, the industry weights corresponding to cross-sector misallocation and adjustment from allocative efficiency should be zero: $\phi_i^{across} = \phi_i^{adjust} = 0$ for all $i \in \mathcal{N}$. Despite this, within-sector price dispersion persists due to pricing frictions, requiring the optimal policy to trade off output gap volatility against within-sector misallocation, as emphasized in the standard New Keynesian literature. Equation in (10) implies that the optimal price-stabilization target is given by $\sum_i \phi_i^* \pi_i = 1 - \prod_i \bar{\mu}_i^{-1}$, and

$$\phi_i^* \equiv \phi_i^{o.g.} + \phi_i^{within}$$

where

$$\phi_i^{o.g.} = 1/\theta_i - 1$$
$$\phi_i^{within} = \frac{\gamma + \varphi}{1 - \prod_{k=1}^N \theta_k}(1/\theta_i - 1)\varepsilon_i \prod_{k=i}^N \theta_k$$

The optimal policy, therefore, allocates larger weights to industries with: (i) higher price stickiness; (ii) larger within-sector elasticities of substitution; and (iii) a more upstream position in the production chain. It is not surprising that our inflation index aligns with the findings of La'O and Tahbaz-Salehi (2022), as there is no cross-sector misallocation in the vertical economy, thus eliminating any supply-side effect of monetary policy.

Example 2. Horizontal Economy

In this section, we use the horizontal economy from section 3.3 as an example to demonstrate how our model primitives characterize the optimal conduct of monetary policy. For simplicity, we assume that the elasticity of substitution in consumption is equal to one.

As illustrated in section 3.3, when initial wedges covary with price rigidities, monetary policy has a supply-side effect. Consequently, the optimal policy component that accounts for the interaction between allocative efficiency and pure technology effect ϕ_i^{adjust} is nonzero. Next, in contrast to the vertical economy, a dispersion on price rigidities result in sectoral relative prices fails to fully reflect their corresponding productivities in response to productivity and monetary shocks, which implies the component of monetary policy targets cross-sector misallocation ϕ_i^{across} is also nonzero. Finally, combining the the components of optimal policy aimed at reducing welfare losses arising from output gap volatility and within-sector price dispersion, the optimal monetary policy in a horizontal economy consists of four components,

$$\phi_i^* \equiv \phi_i^{o.g.} + \phi_i^{within} + \phi_i^{across} + \phi_i^{adjust}$$

where

$$
\begin{aligned}
\phi_i^{o.g.} &= (1/\theta_i - 1)\frac{b_i\bar{\mu}_i^{-1}}{\mathbb{E}_b(\bar{\mu}^{-1})} \\
\phi_i^{within} &= \frac{\gamma+\varphi}{1-\frac{\mathbb{E}_b(\theta\bar{\mu}^{-1})}{\mathbb{E}_b(\bar{\mu}^{-1})}}(1-\theta_i)b_i\varepsilon_i \\
\phi_i^{across} &= (\gamma+\varphi)(1/\theta_i - 1)\frac{\sum_k b_k\bar{\mu}_k^{-1}(\theta_k-\theta_i)}{\mathbb{E}_b(\bar{\mu}^{-1})-\mathbb{E}_b(\theta\bar{\mu}^{-1})}\frac{b_i\bar{\mu}_i^{-1}}{\mathbb{E}_b(\bar{\mu}^{-1})} \\
\phi_i^{adjust} &= -(1+\varphi)(\gamma-1)(1/\theta_i-1)b_i\left(\frac{\bar{\mu}_i^{-1}}{\mathbb{E}_b(\bar{\mu}^{-1})}-1\right)\frac{\text{Cov}_b(\theta,\bar{\mu}^{-1})}{\mathbb{E}_b(\bar{\mu}^{-1})-\mathbb{E}_b(\theta\bar{\mu}^{-1})}.
\end{aligned}
$$

and the optimal inflation bias is determined by

$$
\pi^* = \underbrace{1-\mathbb{E}_b(\bar{\mu}^{-1})}_{\text{Aggregate wedge}} + \underbrace{[1+\varphi+(\gamma-1)\mathbb{E}_b(\bar{\mu}^{-1})]\frac{\text{Cov}_b(\theta,\bar{\mu}^{-1})}{\mathbb{E}_b(\bar{\mu}^{-1})-\mathbb{E}_b(\theta\bar{\mu}^{-1})}}_{\text{Supply-side effect}}
$$

In the horizontal economy, the optimal policy assigns a larger industry weight to industries with (i) higher price stickiness, (ii) larger consumption shares, and (iii) lower initial wedges.[6] It's also noteworthy that when initial wedges are uniform across sectors, the industry weight ϕ_i^* is independent of these initial wedges for all $i \in \mathcal{N}$ and the inflation bias arising from supply-side effect is zero. However, as long as initial markups exist, the aggregate wedge $1 - E_b(\bar{\mu}^{-1})$ still results in an inflation bias.

[6]We can reasonably disregard ϕ_i^{adjust} due to its quantitatively minor significance.