Time-Varying Networks and the Efficacy of Money Without Sticky Prices*

Feng Dong[†] Yi Wen[‡]

This Version: February, 2019

Abstract

We build an analytically tractable model of dynamic production networks with incomplete insurance markets and heterogeneous money demand. We use the model to quantify the classic Baumol-Tobin redistribution channel of monetary policy. Our model can explain (i) the joint distribution of household consumption and money demand and (ii) the strong propagation mechanism of monetary shocks for the business cycle found in empirical VARs across production sectors. We show that the Baumol-Tobin redistribution channel of monetary non-neutrality can be greatly magnified and propagated through endogenous leisure choices and production networks. Our model can account for the hump-shaped impulse responses of sectoral output and employment to monetary shocks, thanks to the endogenous linkage between the distribution of household money demand and firms' input-output coefficient matrix. Our model provides an alternative framework to the Heterogeneous Agent New Keynesian (HANK) model for monetary policy analysis.

Keywords: Dynamic Production Networks, Heterogeneous Money Demand, Redistribution Effect of Money, Wealth Distribution, Time-Varying Velocity of Money, Time-Varying Labor Wedge.

JEL codes: E12, E13, E23, E31, E32, E41, E43, E51.

*Preliminary and comments are welcome. The views expressed here are only those of the individual authors and do not necessarily reflect the official positions of the Federal Reserve Bank of St. Louis, the Federal Reserve System, or the Board of Governors. The other usual disclaim also applies. First Version: February, 2018.

[†]Antai College of Economics and Management, Shanghai Jiao Tong University, Shanghai, China. Tel: (+86) 21-52301590. Email: fengdong@sjtu.edu.cn

[‡]Federal Reserve Bank of St. Louis, P.O. Box 442, St. Louis, MO 63166. Office: (314) 444-8559. Fax: (314) 444-8731. Email: yi.wen@stls.frb.org

1 Introduction

Can monetary shocks significantly affect aggregate resource allocations through channels other than sticky prices? Baumol (1952) and Tobin (1956) were among the first to give a definite answer to the question. They show that money injections can affect aggregate consumption and output through redistribution of the purchasing power of money across heterogeneous agents based on the distribution of money demand. The intuition is that when individuals' demand for liquidity to smooth consumption is heterogeneous, they react to an aggregate monetary injection differently depending on their cash positions, leading to a disproportionally smaller change in the aggregate price level than the injection. For example, due to fixed costs of traveling to banks (or ATM machines), individuals opt to withdraw cash infrequently and hold cash inventories to spread out the fixed costs over time, creating a distribution of real money balances across the population. Hence, an unanticipated aggregate money injection affects only the agents in need of replenishing cash inventories, but not those with sufficient liquidity on hand. Thus, consumers' responses to an aggregate money injection are heterogeneous, leading to a less than proportional increase in the aggregate price level and an increase in aggregate consumption and output. In other words, monetary shocks are not neutral because they affect individual consumption differently by effectively redistributing the purchasing power of money from cash-rich agents to cash-poor agents through an inflation tax, thus resulting in higher aggregate consumption.

Unfortunately, this redistribution channel of monetary non-neutrality in the classic Baumol-Tobin model is too weak to match the empirical magnitude of output responses to monetary shocks (see Section 2 below). An important reason is this: To smooth consumption, rational households opt to carry enough liquidity (cash inventories) as self insurance to buffer both transitory and persistent monetary shocks; consequently, the Baumol-Tobin model lacks both a strong amplification mechanism and a strong propagation mechanism to induce large and persistent movements in aggregate consumption after a money injection. For this very reason, many alternative models with heterogeneous money demand suffer from the same problem, such as the model of Bewley (1980), the heterogeneous-agent cash-in-advance model of Lucas (1980), the quasi-linear preference models of Wen (2010, 2015), and more recently, the production-network model of Dong and Wen (2019).

In particular, Dong and Wen (2019) embed the Bewley model of heterogeneous money demand into the production-network model of Long and Plosser (1983) and show that the multi-sector dynamic stochastic general equilibrium model of Long and Plosser remains analytically tractable under quasi-linear preferences. However, the propagation and amplification mechanism of monetary shocks on aggregate output remains weak despite the fact that an aggregate monetary shock causes also a redistribution effect through production networks due to the non-symmetric property of the input-output table.

In this paper we introduce two modifications into the model of Dong and Wen (2019) to amplify and propagate the redistribution effect of money. The first modification is to replace the idiosyncratic preference shocks in the model of Dong and Wen (2019) by idiosyncratic net-worth shocks. Net-worth shocks are less self-insurable through precautionary money holdings, thus they greatly amplify the redistribution effect of monetary shocks. Interestingly, the Dong-Wen model remains analytically tractable under idiosyncratic net-worth shocks. This property allows Bayesian estimation of model parameters in future works.

The second modification is to introduce "keeping up with the Joneses" (KUWJ) behavior of household leisure choices. Under the KUWJ preferences, higher level of leisure consumption by other agents induces each individual to increase its own leisure consumption; consequently, labor supply has a positive externality and becomes more elastic in responding to monetary shocks. This channel amplifies the non-neutrality of money through the labor-market-wedge channel (see Dong and Wen, 2019) and reinforces the propagation mechanism of production networks.

We calibrate our model to match the input-output table and the joint distribution of household consumption and money holdings in the U.S. economy, and show that the model can quantitatively match the large hump-shaped impulse responses of U.S. aggregate output and sectoral output under monetary shocks.

We interpret the leisure externalities under KUWJ preference as a short cut to search externalities in the labor market. Since more intensive search efforts by others lead to a higher probability of job match for each individual, the labor market is more responsive to aggregate shocks with search frictions than without (see, e.g., Dong, Wang, and Wen, 2018). KUWJ preferences have the same implication as labor-market search but is much simpler to solve.

These two mechanisms are intimately connected and reinforce each other, thanks to the rise of a "labor wedge" in production networks due to market incompleteness. In the original Long-Plosser model with complete markets, the marginal rate of substitution between consumption and leisure always equals the marginal product of labor. Consequently, leisure choices do not affect the input-output coefficients for intermediate goods. However, with incomplete markets, the distribution of money demand endogenously affects household labor supply and the optimal input-output saving ratios of intermediate goods, generating an endogenously time-varying wedge between the marginal rate of substitution and the marginal products of labor across all production sectors. This labor wedge has been shown empirically important for the observed business cycles by Chari, Kehoe, and McGrattan (2007) and Karabarbounis (2014). We theoretically derive this time-varying labor wedge from a network model with heterogenous money demand and show that it serves as a linchpin for production networks and KUWJ preferences to magnify and propagate the Baumol-Tobin redistribution effect.

In addition, introducing idiosyncratic net-worth shocks enables the model to match the empirical joint distribution of household consumption and money demand, whereas the idiosyncratic preference shocks in Dong and Wen (2919) cannot. This success is because iid net-worth shocks are much less self-insurable through real money balances than iid preference shocks. Hence, the Baumol-Tobin redistribution effect is amplified. On the other hand, the existing literature often relies on labor income shocks to generate idiosyncratic risk; but such an approach renders the model intractable and must rely on numerical methods (e.g., see the HANK model of Kaplan, Moll, and Violante, 2018). Since the model's state space is significantly enlarged by multi-sector dynamic production networks, it is unclear how valuable the existing numerical methods are in our model context.

So an additional contribution of our heterogeneous-agent dynamic network (HADA) model is its analytical tractability, as in the original Long-Plosser (1983) model, despite a large state space with heterogeneous capital stocks, heterogeneously household money demand, incomplete markets, and an arbitrarily large number of production sectors. This tractability allows us to analytically study the endogenous linkages among the distribution of money demand, production networks, KUWJ preferences, and the time-varying labor wedge. It also permits using Bayesian method to estimate the model's key parameters in future works.

Literature Review: The work most closely related to ours includes Atalay (2017), Pasten, Schonenle, and Weber (2016) and Ozdagli and Weber (2017). Atalay (2017) develops and estimates a multi-industry model with input-output linkages. His quantitative analysis indicates that industry-specific shocks can account for at least half of aggregate output volatility. Pasten, Schonenle, and Weber (2016) address the propagation of monetary policy shocks in a multi-sector Calvo model with intermediate inputs. Ozdagli and Weber (2017) empirically explore the importance of production networks for the transmission of macroeconomic shocks using stock-market reactions to monetary policy shocks. One of the main differences between this literature and our paper is that we are among the first to study the implications of incomplete financial markets and the time-varying distribution of real money balances for the propagation of monetary shocks through an input-output network structure without the assumption of sticky prices. We also contribute to this literature by offering an analytically tractable network model with heterogenous agents and incomplete markets. Thus, our framework can be readily extended to studying the issues of international trade, optimal capital taxation, and optimal government debts (*a la* Aiyagari, 1994, and others) in a multi-sector setting with money and a realistic input-output structure.

Our paper also relates to the growing literature on production networks. In particular, see Dupor (1999), Gabaix (2011), Foerster, Sarte and Watson (2011), Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi (2012), Acemoglu, Akcigit, and Kerr (2016), Acemoglu, Carvalho, and Tahbaz-Salehi (2017) and Oberfield (2018), among others, for the recent progress on the empirical and theoretical analysis of production networks. For the application of network theory in macro and finance, see Kim and Shin (2012), Kalemli-Ozcan, Kim, Shin, Sørensen, and Yesiltas (2014), Altinoglu (2016), Bigio and La'O (2016), Luo (2017) and Su (2017) for recent contributions on the issue of financial shocks to production chains. Also, see Horvath (1998, 2000) and Shea (2002) for their analyses on sectorial shocks and aggregate fluctuations. Also see Baqaee (2017) and Baqaee and Farhi (2017) for analysis of the macroeconomic impact of microeconomic shocks in a production network. Also see Carvalho (2014), Carvalho and Tahbaz-Salehi (2018), and Foerster, LaRose and Sarte (2018) for literature reviews on production networks.

2 The Model

Time is discrete and proceeds from zero to infinity. The economy consists of heterogenous production sectors and heterogenous households. We describe their problems in the following order.

2.1 Firms

There are $N \geq 1$ production sectors in the set $\mathbf{N}$. Each sector $i \in \mathbf{N}$ has a representative firm that produces output Y_{it} by hiring labor L_{it}, renting capital K_{it} and intermediate goods S_{ijt} from households in centralized markets. Anticipating that factor prices may differ from sector to sector, the problem of a representative firm in sector $i \in \mathbf{N}$ is given by

$$\max_{\{K_{it}, L_{it}, S_{ijt}\}} Q_{it} \left\{ Y_{it} - R^k_{it} K_{it} - W_{it} L_{it} - \sum_{j=1}^{N} R^s_{ijt} S_{ijt} \right\}, \tag{1}$$

subject to the production technology

$$Y_{it} = A_{it} K_{it}^{\alpha_i^k} L_{it}^{\alpha_i^l} \left(\prod_{j=1}^{N} S_{ijt}^{\omega_{ij}} \right)^{\alpha_i^s}, \text{ for } i \in \mathbf{N}, \tag{2}$$

where Q_{it} is sector i's relative price in terms of aggregate value added, R^k_{it} the sector's rental price of capital, W_{it} the sector's real wage, and R^s_{ijt} the sector's purchase price of intermediate good produced by sector $j \in \mathbf{N}$, and A_{it} a sector-specific productivity shock. The production technology in each sector is constant returns to scale with $\sum_{j=1}^{N} \omega_{ij} = 1$ and $\alpha_i^k + \alpha_i^l + \alpha_i^s = 1$ for all $i \in \mathbf{N}$.

2.2 Households

Households supply labor, save and rent out capital and intermediate goods to firms. Households also hold money as a store of value to smooth consumption and self-insure against idiosyncratic risk, as in Bewley (1980).

Specifically, there is a continuum of *ex ante* identical households indexed by $\iota \in [0,1]$. Each household is subject to an idiosyncratic shock $\theta_t(\iota)$ to its net-worth $x_t(\iota)$, θ_t has the distribution $F(\theta) \equiv \Pr[\theta(\iota) \leq \theta]$ with support $[\theta_{\min}, \theta_{\max}]$. Leisure enters the utility function linearly as in Wen (2015). Each household chooses a $N \times 1$ consumption vector $\mathbf{c}_t(\iota) = \{c_{jt}(\iota)\}_{j \in \mathbf{N}}$, a $N \times 1$ labor supply vector $\mathbf{l}_t(\iota) = \{l_{jt}(\iota)\}_{j \in \mathbf{N}}$, the nominal money balances $m_{t+1}(\iota)$, a $N \times N$ matrix of savings of intermediate inputs $\mathbf{s}_{t+1}(\iota) = \{s_{ij,t+1}(\iota)\}_{i,j \in \mathbf{N}}$, and a $N \times N$ matrix of savings of capital investment inputs $\mathbf{z}_t(\iota) = \{z_{ijt}(\iota)\}_{i,j \in \mathbf{N}}$ to maximize lifetime utility.

We assume "keeping-up-with-Joneses" (KUWJ) preferences on leisure consumption, in that the aggregate level of labor supply L_t reduces the marginal-utility cost of each individual's labor supply but is taken as given by individual households. The period utility function is given by:

$$u(\mathbf{c}_t(\iota), \mathbf{l}_t(\iota)) = \sum_{j=1}^{N} \varphi_j \log c_{jt}(\iota) - \psi \frac{\sum_{j=1}^{N} l_{jt}(\iota)}{(L_t/L)^{\nu}}, \tag{3}$$

where the parameters satisfy $\sum_{j=1}^{N} \varphi_j = 1$, with $\varphi_j > 0$ for all $j \in N$;[1] $L_t = \sum_{j=1}^{N} L_{jt} = \sum_{j=1}^{N} \int_0^1 l_{jt}(\iota)\, d\iota$ is the level of aggregate labor supply taken as given by the household, with L as its steady-state value; $\nu \geq 0$ measures the degree of negative externality from other people's leisure choices on an individual's disutility of working—the KUWJ preferences. For simplicity, we introduce externality at the aggregate level.[2] Hence, a higher (than steady-state) level of leisure choice by everyone else reduces household ι's marginal utility of leisure, or working harder by others induces each individual to work harder as well, as if there is a search friction in the labor market. The channel of KUWJ can be shut down by setting $\nu = 0$.

Capital is sector-specific. As firm owners, households save by holding money and renting capital and intermediate goods to firms. To ensure tractability of the model with capital, we assume that the law of motion of capital accumulation facing household ι in each sector j is given by the log-linear form:

$$k_{j,t+1}(\iota) = (k_{jt}(\iota))^{1-\delta_j} (i_{jt}(\iota)/\delta_j)^{\delta_j}, \tag{5}$$

[1] Similar to Bigio and La'O (2017), we can interpret the basket of consumption goods as a composite consumption good such that $u(c,l) = \theta \log c - \psi l$, where $c \equiv \prod_{j=1}^{N} c_j^{\varphi_j}$.

[2] Alternatively, we can model the externality at the sectoral level. Then the utility in equation (3) becomes

$$u(\mathbf{c}(\iota)_t, \mathbf{l}(\iota)_t) = \sum_{j=1}^{N} \varphi_j \log c_{jt}(\iota) - \psi \sum_{j=1}^{N} \frac{l_{jt}(\iota)}{(L_{jt}/L_j)^{\nu_j}}. \tag{4}$$

where $\delta_j \in (0,1)$ denotes the sector-specific depreciation rate of capital and i_{jt} denotes household investment in sector j's capital stock. In the steady state, we have $i_j(\iota) = \delta_j k_j(\iota)$. Note that if $\delta_j = 1$, then $k_{j,t+1}(\iota) = i_{jt}(\iota)$. As will become clear later, the log-linear law of motion of capital in equation (5) guarantees tractability of network models with N dimensional durable capital stocks.

As in Foerster, Sarte, and Watson (2011), we also assume that investment good in each sector j is formed by combining outputs from all production sectors according to the CRS technology:

$$i_{jt}(\iota) = \epsilon_{jt} \prod_{i=1}^{N} (z_{jit}(\iota))^{\phi_{ji}}, \tag{6}$$

where $\sum_{i=1}^{N} \phi_{ji} = 1$, $\phi_{ji} > 0$, and ϵ_{jt} denotes sector-specific investment technology shock.

Define household ι's real income from sector j as

$$\tilde{y}_{jt}(\iota) \equiv R_{jt}^{k} k_{jt}(\iota) + \sum_{i=1}^{N} R_{jit}^{s} s_{jit}(\iota) + W_{jt} l_{jt}(\iota), \tag{7}$$

and define the household's "cash on hand" or "net-worth" as

$$x_t(\iota) \equiv \frac{m_t + \tau_t}{P_t} + \sum_{j=1}^{N} Q_{jt} \tilde{y}_{jt}(\iota) - \sum_{j=1}^{N}\sum_{i=1}^{N} Q_{jt} s_{ij,t+1}(\iota) - \sum_{j=1}^{N}\sum_{i=1}^{N} Q_{jt} z_{ijt}(\iota) - T_t, \tag{8}$$

then the household's budget constraint is given by

$$\sum_{j=1}^{N} Q_{jt} c_{jt}(\iota) + \frac{m_{t+1}(\iota)}{P_t} \leq [\varepsilon + \theta_t(\iota)] \cdot x_t(\iota), \tag{9}$$

where $\frac{m_{t+1}(\iota)}{P_t}$ denotes real money demand chosen in period t, P_t denotes the nominal price of aggregate output, τ_t denotes a random money injection that is equally distributed across households, and $T_t = \sum Q_{jt} T_{jt}$ denotes a lump-sum tax that is identically levied across households.[3]

Note that each household's net-worth is subject to an *iid* idiosyncratic shock $\theta_t(\iota)$, which is the only source of idiosyncratic uncertainty in the model; so that the net-worth on the right hand side of equation (9) has a multiplier $\varepsilon + \theta_t(\iota)$, where $\varepsilon \in (0,1)$ is a constant, and $\theta_t(\iota)$ is the idiosyncratic shock to household net-worth. We normalize the mean of $\theta_t(\iota)$ with $\mathbb{E}(\theta_t(\iota)) = 1 - \varepsilon$, so that the average value of the multiplier $\varepsilon + \theta_t(\iota)$ equals 1.

This normalization suggests that idiosyncratic net-worth shocks do not cause distortion to household resource constraint at the aggregate level. The idiosyncratic net-worth shock

[3] As will be shown later, log-linear preferences may imply negative labor supply in some states if the household's real money balances are too high, so we introduce a lump sum tax to ensure positive labor supply across all households in each production sector. This interior solution for labor supply is critical for closed-form solutions in our heterogeneous-agent economy.

is thus analogous to an idiosyncratic redistributionary "tax" shock, as it implies that nature redistributes a portion of aggregate household's net wealth randomly across households without changing the level of aggregate net-worth. Since the shocks are assumed to be *iid*, there is no correlation in household net worth across agents.

Then the problem of household ι is to solve

$$\max \mathbb{E}_0 \left\{ \sum_{t=0}^{\infty} \beta^t \sum_{j=1}^{N} \varphi_j \ln c_{jt}(\iota) - \psi_t \sum_{j=1}^{N} l_{jt}(\iota) \right\}, \tag{10}$$

subject to the budget constraint (9) and the following no-short-sale (borrowing) constraint on nominal balances:

$$m(\iota)_{t+1} \geq 0; \tag{11}$$

where $\psi_t \equiv \frac{\psi}{(L_t/L)^{\nu}}$ in the utility function denotes the time-varying marginal cost (disutility) of work.

The information structure of the household problem is specified as follows. In each period t, a household must determine its net-worth x_t before observing the net-worth shock θ_t, and then choose consumption and money holdings after observing θ_t. Consequently, the idiosyncratic shock gives rise to incentives for households to hold money as a self-insurance device under the borrowing constraint (11).

Unlike idiosyncratic labor income shocks commonly specified in standard Bewley (1980) or Aiyagari (1994) models, net-worth shocks are less insurable through precautionary savings even when the shocks are *iid.* This not only helps the model (among others) to be analytically tractable but also allows it to match the highly unequal distributions of money demand in the U.S. household data. A realistic implication of the distribution of money holdings is important to gage the Baumol-Tobin redistribution effect.

Analytical tractability of the model on the household side also requires other important features. Specifically, following Wen (2015), we assume log-linear preferences and a special information structure by dividing each period t into two sub-periods: decisions for labor supply, capital investment, and savings for intermediate goods are made in the first subperiod before $\theta_t(\iota)$ is realized; while decisions for consumption and money holdings are determined in the second subperiod after observing $\theta_t(\iota)$. Consequently, all components in net-worth $x_t(\iota)$ (except aggregate state variables) are predetermined with respect to θ_t in each period (i.e., determined before the realization of $\theta_t(\iota)$).

Since money is not required (or imposed from outside) as a medium of exchange—unlike the cash-in-advance model, choosing $m_t(\iota) = 0$ by all households for all periods t is always an equilibrium. In the remaining part of the paper, we focus on the case where money is valued in equilibrium and is thus "essential," i.e., the equilibrium aggregate nominal price $P_t < \infty$ for all $t \geq 0$.

Market Clearing Conditions: The market clearing conditions for all types of goods (investment and intermediate goods) and labor in each sector j are given, respectively, by

$$C_{jt} + \sum_{i=1}^{N} Z_{ijt} + \sum_{i=1}^{N} S_{ij,t+1} + T_{jt} = Y_{jt}, \tag{12}$$

$$L_{jt} = \int_0^1 l_{jt}(\iota)\, d\iota. \tag{13}$$

To simplify notations further, in what follows we suppress the household index ι unless confusion may arise.

2.3 Characterization of Household Decision Rules

Proposition 1 *The household decision rules follow a cutoff strategy. Denoting θ_t^* as the cutoff of net-worth shocks and $W_t = Q_{jt}W_{jt}$ as the aggregate competitive wage rate across sectors (under perfect labor mobility), then given the sequences of prices $\left\{W_{jt}, R_{it}^k, R_{ijt}^s, Q_{jt}\right\}_{i,j\in\mathbf{N}}$, the policy functions of optimal target of cash on hand $(\varepsilon + \theta_t^*)\, x_t$, consumption c_{jt}, and real money demand $\frac{m_{t+1}}{P_t}$ can be analytically characterized, respectively, by:*

$$(\varepsilon + \theta_t^*)\, x_t = \frac{W_t \Gamma(\theta_t^*)}{\psi_t}, \tag{14}$$

$$c_{jt} = \frac{\varphi_j}{Q_{jt}} \min\left\{1, \frac{\varepsilon + \theta_t}{\varepsilon + \theta_t^*}\right\} (\varepsilon + \theta_t^*)\, x_t, \tag{15}$$

$$\frac{m_{t+1}}{P_t} = \max\left\{\frac{\theta_t(\iota) - \theta_t^*}{\varepsilon + \theta_t^*}, 0\right\} (\varepsilon + \theta_t^*)\, x_t, \tag{16}$$

Proof: See Appendix. ■

Proposition 2 *The cutoff θ_t^* is determined by the following Euler equation,*

$$\frac{\psi_t}{W_t} = \beta \mathbb{E}_t \frac{P_t}{P_{t+1}} \frac{\psi_{t+1}}{W_{t+1}} \Gamma(\theta_t^*), \tag{17}$$

in which the function $\Gamma(\theta_t^*)$ captures the liquidity premium of money and it is given by

$$\Gamma(\theta_t^*) \equiv \int \max\{\varepsilon + \theta, \varepsilon + \theta^*\}\, dF(\theta) \geq 1. \tag{18}$$

Proof: See Appendix. ■

Proposition 3 *Furthermore, optimal savings of intermediate goods $s_{ij,t+1}$ and investment inputs $z_{ij,t}$ are given, respectively, by the following intertemporal Euler equations:*

$$Q_{jt}\frac{\psi_t}{W_t} = \beta \mathbb{E}_t Q_{i,t+1} \frac{\psi_{t+1}}{W_{t+1}} R_{ij,t+1}^s, \tag{19}$$

$$\lambda_{jt} = \beta \mathbb{E}_t \mu_{t+1} Q_{j,t+1} R^k_{j,t+1} + \beta \mathbb{E}_t \lambda_{j,t+1} \frac{\partial k_{j,t+2}}{\partial k_{j,t+1}}, \tag{20}$$

where (λ_{jt}, μ_t) *denote, respectively, the Lagrangian multipliers of the of motion for capital (5) and the household budget constraint (9).*

Proof: See Appendix. ■

The Liquidity Demand Theory of Money. Proposition 2 shows that the cutoff θ_t^* is independent of the history of individual household's net-worth shocks, $\{\theta_0, \theta_1, ..., \theta_t\}$, but depends only on the aggregate state of the economy. Consequently, the net-worth x_t and cash on hand $(\varepsilon + \theta_t) x_t$ are also independent of the history of household wealth shocks, as revealed in Proposition 1 by equation (14). Therefore, the cutoff θ_t^* is an aggregate variable and serves as a sufficient statistic to fully characterize the distribution of household money demand and consumption.

In particular, Proposition 1 shows that given a distribution of the effective net-worth $(\varepsilon + \theta_t) x_t$, there exists a target level of cash on hand,

$$x_t^* \equiv (\varepsilon + \theta_t^*) x_t, \tag{21}$$

such that consumption of good j is proportional to the target level of cash on hand in equation (15), with the marginal propensity to consume governed by the concave function $\min\left\{1, \frac{\varepsilon+\theta}{\varepsilon+\theta^*}\right\} \leq 1$; implying that when the net-worth shock is below the cutoff $(\theta < \theta^*)$, consumption is only a fraction$\frac{(\varepsilon+\theta)}{(\varepsilon+\theta^*)} < 1$ of the target x^*, and when the net-worth shock is above the cutoff $(\theta > \theta^*)$, consumption equals the target x^*.

Complementing this consumption behavior, equation (16) says that household real money demand $\frac{m_{t+1}}{P_t}$ is a buffer stock: It is zero if cash on hand or net-worth is temporarily below target ($\theta < \theta^*$ or $(\varepsilon + \theta) x < x^*$), and is strictly positive if net-worth is temporarily above target ($\theta \geq \theta^*$ or $(\varepsilon + \theta) \geq x^*$). This property suggests that money serves as a precautionary store of value—it provides self-insurance in case the future cash on hand is below target.

In other words, the optimal level of money holdings is exactly like optimal inventories (Wen, 2011)—it determines the probability of a "cash stockout", depending on the cutoff θ_t^*. Such precautionary saving behavior under borrowing constraints is also analyzed by Deaton (1991) even though Deaton is unable to show these properties analytically.

The intuition is that each household can set labor income (in advance of the realization of θ_t) to target an optimal level of net-worth or cash on hand, so that x_t (or x_t^*) is optimal *ex anti* with respect to the distribution of θ. This in turn implies that regardless of the initial value of real money balances $\frac{m_t}{P_t}$, the household always adjusts labor supply (wage income) to ensure that its cash on hand is optimal (on target) to buffer net-worth shocks and to smooth consumption. Given that θ is iid and the marginal cost of leisure is constant, all households opt to choose the same target level of x (or x^*) regardless of their initial money balances carried over from last period.

Notice that the first-order condition for labor choices yields $\psi_t/W_t = \int (\varepsilon + \theta_t) \mu_t dF(\theta)$, where μ is the Lagrangian multiplier for the household budget constraint in equation (9). Also, the first-order condition for household consumption implies $1/\sum_{j=1}^{N} Q_{jt} c_{jt}(\iota) = \mu_t(\iota)$. Thus, the aggregate real wage equals the average marginal utilities of consumption across households. Denoting $\Lambda_t = \int (\varepsilon + \theta_t) \mu_t dF(\theta) = \psi_t/W_t$ as the average marginal utilities of household consumption, then equation (17) can be rewritten in a more conventional form as

$$1 = \beta \mathbb{E}_t \frac{\Lambda_{t+1}}{\Lambda_t} \frac{P_t}{P_{t+1}} \Gamma(\theta_t^*), \tag{22}$$

which indicates intertemporal trade-off of holding money, where $\beta\Lambda_{t+1}/\Lambda_t$ pertains to the average stochastic discount factor or the average pricing kernel; and P_t/P_{t+1} is the inverse of the inflation rate. Hence, the expected rate of return to money is given by the discounted inflation-adjusted liquidity premium, $\Gamma(\theta^*) = \int \max(\varepsilon + \theta, \varepsilon + \theta^*) dF(\theta) > 1$ (since $\theta^* > \theta_{\min}$ and $E(\varepsilon + \theta) = 1$), which implies that the option value of one dollar exceeds 1 because it provides liquidity in the case of a low net-worth shock. This is why money has positive value in equilibrium despite the fact that its real rate of return is negative and dominated by the rate of time preference $1/\beta$ and interest bearing assets.

The optimal level of cash holdings (money demand) is such that the probability of running out of cash (i.e., being liquidity constrained) is strictly positive (i.e., $F(\theta_t^*) > 0$) unless the cost of holding money is zero (such as under the Friedman rule). Also note that aggregate shocks will affect the distribution of money holdings across households by affecting the cutoff θ_t^*.

The liquidity premium $\Gamma(\theta^*)$ is increasing in the cutoff θ^*, $\frac{\partial\Gamma(\theta^*)}{\partial\theta^*} > 0$, because a higher cutoff implies a higher probability of running out of cash. So, too high a level of cash on hand implies excessively low probability of a binding liquidity constraint, which is too costly given the inflation tax. On the other hand, too low a level of cash on hand implies a high probability of being cash constrained *ex post*. Hence, the optimal target level of cash on hand x is chosen by adjusting labor supply according to the distribution of θ such that x is the same across households in each period regardless of individual history and initial money balances. This optimal choice of cash on hand simultaneously determines the optimal cutoff θ_t^*.

Equation (19) in Proposition (3) is the Euler equation for household's demand for intermediate input $s_{ij,t+1}$, which can be rewritten as

$$Q_{jt} = \beta \mathbb{E}_t \frac{\Lambda_{t+1}}{\Lambda_t} Q_{i,t+1} R_{ij,t+1}^s. \tag{23}$$

The LHS of the above equation is the cost of purchasing one unit of intermediate input $s_{ij,t+1}$ in period t, and the RHS is the next period. Similarly, equation (20) is essentially the demand function for investment input z_{ijt}, where the LHS and RHS represent the marginal cost and benefit of z_{ijt}, respectively.

Finally, assuming a sufficiently large time endowment $\bar{l}$ and lump-sum tax T—to ensure an interior solution for labor supply across households in all sectors, then it must be true that real wages are equalized across sectors: $Q_{jt}W_{jt} = W_t$ for all $j \in N$. Denoting $l_t(\iota) \equiv \sum_{j=1}^{N} l_{jt}(\iota)$, then combining equations (9) and the condition on aggregate wage gives the labor income of household ι as

$$W_t l_t(\iota) = x_t(\iota) + T_t - \frac{m_t(\iota) + \tau_t}{P_t} - \sum_{j=1}^{N} Q_{jt}\left(R_{jt}^k k_{jt}(\iota) + \sum_{i=1}^{N} R_{jit}^s s_{jit}(\iota) - s_{ij,t+1}(\iota) - z_{ijt}(\iota)\right). \tag{24}$$

2.4 Equilibrium Analysis

2.4.1 Aggregation across Households.

As mentioned earlier, an important property of the model is that the cutoff θ_t^* and the optimal net worth x_t (hence the target level of cash on hand $x^* \equiv (\varepsilon + \theta_t^*) x_t$) are independent of household history but depend only on the aggregate state. This important property in conjunction with the closed-form decision rules in Proposition 1 permit aggregation by the law of large numbers in our model, so that the aggregate variables in the model form a dynamic system of non-linear equations that are virtually indifferent from those of any standard RBC models, thanks to the fact that the distribution of households is fully captured by the cutoff θ_t^*, which is itself analytically tractable according to equation (17).

For convenience, we use upper-case letters to denote aggregate variables across households (i.e., $Z_{ijt} \equiv \int z_{ijt}(\iota) dF(\theta), X_t = x_t$) and bold letters to denote the matrix-form or vector-form of these variables (i.e., $\mathbf{Z}_t \equiv \{Z_{ijt}\}_{ij\in\mathbf{N}}$) as well as aggregate prices. Then, integrating individual policy functions in Proposition 1 under the law of large numbers, we can obtain the dynamic system of equations that govern the path of $\{\mathbf{Q}_t$, $\mathbf{R}_t^k$, $\mathbf{R}_t^s$, $\mathbf{W}_t$, $\mathbf{L}_t$, $\mathbf{C}_t$, $\mathbf{S}_t$, $\mathbf{Z}_t$, $\mathbf{I}_t$, $\mathbf{Y}_t$, X_t, θ_t^*, W_t, Λ_t, ψ_t, M_{t+1}, $P_t\}$ in a competitive equilibrium.

Proposition 4 *Denoting $D(\theta^*) \equiv \int \min(\varepsilon + \theta, \varepsilon + \theta^*) dF(\theta)$ and $H(\theta^*) \equiv 1 - D(\theta^*)$, the dynamic system of equations to solve for $\{\mathbf{Q}_t$, $\mathbf{R}_t^k$, $\mathbf{R}_t^s$, $\mathbf{W}_t$, $\mathbf{L}_t$, $\mathbf{C}_t$, $\mathbf{S}_t$, $\mathbf{Z}_t$, $\mathbf{I}_t$, $\mathbf{Y}_t$, X_t, θ_t^*,W_t, Λ_t, ψ_t, M_{t+1}, $P_t\}$ are characterized by the following equations:*

$$C_{jt} = \frac{\varphi_j}{Q_{jt}} D(\theta_t^*) X_t, \tag{25}$$

$$\frac{M_{t+1}}{P_t} = H(\theta_t^*) X_t, \tag{26}$$

$$\frac{1}{P_t} = \beta \mathbb{E}_t \frac{\Lambda_{t+1}}{\Lambda_t} \frac{1}{P_{t+1}} \Gamma(\theta_t^*), \tag{27}$$

$$Q_{jt} = \beta \mathbb{E}_t \frac{\Lambda_{t+1}}{\Lambda_t} Q_{i,t+1} \alpha_i^s \omega_{ij} \frac{Y_{i,t+1}}{S_{ij,t+1}}, \tag{28}$$

$$Q_{jt} = \beta \mathbb{E}_t \frac{\Lambda_{t+1}}{\Lambda_t} Q_{i,t+1} \delta_i \alpha_i^k \phi_{ij} \frac{Y_{i,t+1}}{Z_{ijt}} + \beta \mathbb{E}_t \frac{\Lambda_{t+1}}{\Lambda_t} Q_{j,t+1} (1-\delta_i) \frac{Z_{ij,t+1}}{Z_{ijt}}, \tag{29}$$

$$(\varepsilon + \theta_t^*) X_t = \frac{\Gamma(\theta_t^*)}{\Lambda_t}, \tag{30}$$

$$Y_{it} = A_{it} K_{it}^{\alpha_i^k} L_{it}^{\alpha_i^l} \left(\prod_{j=1}^{N} S_{ijt}^{\omega_{ij}} \right)^{\alpha_i^s}, \tag{31}$$

$$K_{i,t+1} = K_{it}^{1-\delta_i} (I_{it}/\delta_i)^{\delta_i}, \tag{32}$$

$$I_{it} = \epsilon_{it} \prod_{j=1}^{N} Z_{ijt}^{\phi_{ij}}, \tag{33}$$

$$R_{jt}^k = \alpha_j^k \frac{Y_{jt}}{K_{jt}}, \tag{34}$$

$$W_{jt} = \alpha_j^l \frac{Y_{jt}}{L_{jt}}, \tag{35}$$

$$R_{ijt}^s = \alpha_i^s \omega_{ij} \frac{Y_{it}}{S_{ijt}}, \tag{36}$$

$$W_t = Q_{jt} W_{jt}, \tag{37}$$

$$C_{jt} + \sum_{i=1}^{N} S_{ij,t+1} + \sum_{i=1}^{N} Z_{ijt} + T_{jt} = Y_{jt}, \tag{38}$$

$$X_t = \frac{M_t + \tau_t}{P_t} + \sum_{j=1}^{N} Q_{jt} \left[R_{jt}^k K_{jt} + W_{jt} L_{jt} + \sum_{i=1}^{N} \left(R_{jit}^s S_{jit} - S_{ij,t+1} - Z_{ijt} \right) \right] - T_t, \tag{39}$$

$$\overline{M}_{t+1} = \overline{M} + \overline{\tau}_t, \tag{40}$$

$$\psi_t = \frac{\psi}{(L_t/L)^{\nu}}, \tag{41}$$

$$\Lambda_t = \frac{\psi_t}{W_t}, \tag{42}$$

for $\{i, j\} \in \mathbf{N}$,*where* T_{jt} *denotes sectoral lump-sum tax or government spending taken as given by all agents, and* $\bar{M}$ *denotes aggregate money supply.*

Proof: See Appendix. ■

Equation (25) and (26) represent, respectively, consumption for sectoral good j and the aggregate demand for real money balances. Since $\mathbb{E}(\theta) = 1 - \varepsilon$, the average marginal propensity to consume is given by

$$D(\theta^*) = \varepsilon + \int_{\theta < \theta^*} \theta dF + \int_{\theta \geq \theta^*} \theta^* dF \in (\varepsilon + \theta_{\min}, 1), \tag{43}$$

and the average marginal propensity to holding money is given by

$$H(\theta^*) \equiv 1 - D(\theta^*) = \int_{\theta \geq \theta^*} [\theta - \theta^*] \, dF \in (0, 1 - \varepsilon - \theta_{\min}). \tag{44}$$

In equation (40) we assume that the supply of money $\overline{M}_t$ (aggregate money stock) is stationary around the mean $\overline{M}$ and the monetary injection shock $\bar{\tau}_t$ follows an AR(1) process:

$$\bar{\tau}_t = \rho_\tau \bar{\tau}_{t-1} + \varepsilon_t^\tau. \tag{45}$$

Such a specification is consistent with our empirical VAR analysis in a later section where the stock of money is detrended; it implies that any injected money is eventually taken out of the economy, so there is no permanent effect on the inflation rate, as in the U.S. qualitative easing (QE) episodes after the recent financial crisis.[4]

2.4.2 Consumption Velocity of Money

Define the aggregate consumption across both households and goods sectors as

$$C_t = \sum_{j=1}^{N} Q_{jt} C_{jt}. \tag{46}$$

Then, with the normalization $\sum_j \varphi_j = 1$, equation (25) immediately implies

$$C_t = D(\theta_t^*) X_t. \tag{47}$$

Combining equations (26), (47), and the money-market clearing condition (**??**) yields

$$P_t C_t = M_{t+1} \frac{D(\theta_t^*)}{H(\theta_t^*)}.$$

Then the aggregate (consumption) velocity of money is given by

$$V(\theta_t^*) \equiv \frac{P_t C_t}{M_{t+1}} = \frac{D(\theta_t^*)}{H(\theta_t^*)} \in (0, \infty). \tag{48}$$

Thus, money velocity in our model is a function only of the *distribution of money demand* (the cutoff θ^*) across households, which in turn depends on the aggregate state variables, including monetary shocks. This time-varying velocity of money with the support $\left[\frac{\varepsilon+\theta_{\min}}{1-(\varepsilon+\theta_{\min})}, \infty\right)$ is in sharp contrast to the representative-agent cash-in-advance (CIA) models that typically implies a constant velocity of 1.[5] Our model can also match the variability of money velocity in the data.

[4] Since aggregate money demand follows the law of motion, $M_{t+1} = M_t + \tau_t$, then the money market clearing condition, $M_{t+1} = \bar{M}_{t+1}$, implies that cash received by households in each period is given by $\tau_t = \bar{\tau}_t - \bar{\tau}_{t-1}$, which has an ARMA(1,1) representation: $\tau_t = \rho_\tau \tau_{t-1} + \varepsilon_t^\tau - \varepsilon_{t-1}^\tau$, suggesting that aggregate money demand is also stationary instead of following a random walk.

[5] Since $D_t = \mathbb{E}_t \min(\varepsilon + \theta, \varepsilon + \theta_t^*) \in [\varepsilon + \theta_{\min}, 1]$, and since $V_t \equiv \frac{D_t}{1-D_t}$ increases with D_t, easily we know that $V_t \in \left[\frac{\varepsilon+\theta_{\min}}{1-(\varepsilon+\theta_{\min})}, \infty\right)$.

Remark 1 *Following Jones (2013) and Bigio and La'O (2016), the composite consumption bundle of an individual household can be defined as* $c(\iota)_t = \prod_{j=1}^{N} \left(c(\iota)_{jt}\right)^{\varphi_j}$. *Then we can define aggregate consumption as*

$$C_t \equiv \int_0^1 c_t(\iota)\, d\iota = \int c_t(\iota)\, dF(\theta) = \frac{D(\theta_t^*) X_t}{Q_t}, \tag{49}$$

where the price index Q_t *is defined by*

$$Q_t \equiv \prod_{j=1}^{N} \left(\frac{Q_{jt}}{\varphi_j}\right)^{\varphi_j}. \tag{50}$$

By the normalization $Q_t = 1$, *we have* $C_t = D(\theta_t^*)X_t$, *which coincides with equation (47).*

Thus, the aggregate marginal propensity to consume from aggregate net worth X_t is still given by $D(\theta_t^*)$, and the aggregate marginal propensity to save in the form of real money balances is still $H(\theta_t^*)$. These marginal propensities are time varying purely because the distribution of household money demand (characterized by θ_t^*) is time varying. This is an critical departure from the complete-market RBC model of Long and Plosser (1983).

2.4.3 Time-Varying Input-Output Coefficients

Our model is reduced to that of Dupor (1999) if we (i) remove the incomplete-market environment by letting the variance of idiosyncratic shocks $Var(\theta) \to 0$; and (ii) assume full depreciation of capital, i.e., $\delta_j = 1$ for all sector $j \in \mathbf{N}$. In addition, if we further remove capital input from the production function by setting $\alpha_j^k = 0$ for all $j \in \mathbf{N}$, then the model is further reduced to the original model of Long and Plosser (1983).

Interestingly, despite incomplete markets and durable capital stocks in all sectors of production, we can still analytically derive closed-form solutions and express sectoral consumption C_{jt} as a linear function of sectoral output Y_{jt}, as in Long and Plosser (1983), except the coefficients are now *time-varying*, as the following Proposition shows.

Proposition 5 *The optimal consumption and savings of investment goods and intermediate goods in sector* j *are all proportional to their respective sectoral output:*

$$C_{jt} = \xi_{jt}^c Y_{jt} \equiv \frac{\varphi_j}{\gamma_{jt}} Y_{jt}, \tag{51}$$

$$S_{ij,t+1} = \xi_{ijt}^s Y_{jt} \equiv \left[\beta \alpha_i^s \frac{\gamma_{it}}{\gamma_{jt}}\right] \widetilde{\omega}_{ijt} Y_{jt}, \tag{52}$$

$$Z_{ijt} = \xi_{ijt}^z Y_{jt}, \tag{53}$$

where the time-varying coefficient $\widetilde{\omega}_{ijt} \equiv \omega_{ij}\mathbb{E}_t\left[\frac{\psi_{t+1}}{\psi_t}\frac{L_{i,t+1}}{L_{it}}\right]$ *will enter the model-implied IO table and the time-varying coefficients* $\left\{\gamma_{jt}, \xi_{ijt}^z\right\}$ *are jointly determined by the following two* *dynamic equations:*

$$\left(1-g_{jt}-\sum_{i=1}^{N}\xi_{ijt}^{z}\right)\gamma_{jt}=\varphi_j+\beta\sum_{i=1}^{N}\gamma_{it}\alpha_i^s\omega_{ij}\mathbb{E}_t\left(\frac{\psi_{t+1}L_{i,t+1}}{\psi_t L_{it}}\right), \tag{54}$$

$$\xi_{ijt}^{z}\gamma_{jt}=\beta\mathbb{E}_t\left\{\frac{\psi_{t+1}}{\psi_t}\left[\frac{L_{i,t+1}}{L_{it}}\delta_i\phi_{ij}\alpha_i^k\gamma_{it}+\frac{L_{j,t+1}}{L_{jt}}\left(1-\delta_i\right)\xi_{ij,t+1}^{z}\gamma_{jt}\right]\right\} \tag{55}$$

where for simplicity and without loss of generality we have assumed that $g_{jt}\equiv T_{jt}/Y_{jt}$*, which denotes the lump-sum tax to output ratio in sector* j *and is taken as given by all agents. These two dynamic equations can be expressed more compactly in matrix form:*

$$\boldsymbol{\gamma}_t=\left(\mathbf{E}-\beta\widehat{\boldsymbol{\Omega}}_t'\right)^{-1}\boldsymbol{\varphi}, \tag{56}$$

$$\widehat{\boldsymbol{\Omega}}_t\equiv\frac{1}{\beta}\mathbf{Diag}\left(\mathbf{g}_t+\widetilde{\boldsymbol{\xi}}_t^{z}\right)+\mathbf{Diag}\left(\boldsymbol{\alpha}^{\mathbf{s}}\right)\widetilde{\boldsymbol{\Omega}}_t, \tag{57}$$

where $\boldsymbol{\gamma}_t$ *is a vector with elements* γ_{jt}*,* $\mathbf{E}$ *is an identity matrix,* $\boldsymbol{\varphi}$ *is a vector with elements* φ_j*,* $\mathbf{g}_t$ *is a matrix with column vector* g_{jt}*,* $\widetilde{\boldsymbol{\xi}}_t^{z}$ *is a matrix with column vector* $\widetilde{\xi}_{jt}^{z}\equiv\sum_{i=1}^{N}\xi_{ijt}^{z}$*,* $\boldsymbol{\alpha}^s$ *is a matrix with column vector* α_j^s*, and* $\widetilde{\boldsymbol{\Omega}}_t$ *is the input-output matrix with elements* $\widetilde{\omega}_{ijt}\equiv\omega_{ij}\mathbb{E}_t\left[\frac{\psi_{t+1}}{\psi_t}\frac{L_{i,t+1}}{L_{it}}\right]$.

Proof: See Appendix. ■

Notice that in the absence of (i) capital ($\alpha_j^k=0$ for all $j\in\mathbf{N}$), (ii) of KUWJ preferences ($\psi_t=1$), (iii) of lump-sum taxies ($T_{jt}=0$), and (iv) of idiosyncratic shocks ($Var\left(\theta_t\right)=0$), we then have $\mathbf{g}_t=\tilde{\boldsymbol{\xi}}_t^{z}=0$, $\frac{\psi_{t+1}}{\psi_t}\frac{L_{i,t+1}}{L_{it}}=1$, and $\tilde{\omega}_{ijt}=\omega_{ij}$, so equations (**??**) and (57) become

$$\boldsymbol{\gamma}=\left(\mathbf{E}-\beta\widehat{\boldsymbol{\Omega}}'\right)^{-1}\boldsymbol{\varphi}=\left(\mathbf{E}-\boldsymbol{\beta}\mathbf{Diag}\left(\boldsymbol{\alpha}^{\mathbf{s}}\right)\boldsymbol{\Omega}'\right)^{-\mathbf{1}}\boldsymbol{\varphi}, \tag{58}$$

where $\boldsymbol{\Omega}=\{\omega_{ij}\}_{ij\in\mathbf{N}}$ is the input-output coefficient matrix in the original Long-Plosser model. Accordingly, the consumption and saving equations in (51) and (52) are reduced to their counterparts in the original Long-Plosser model:

$$C_{jt}=\frac{\varphi_j}{\gamma_j}Y_{jt}, \tag{59}$$

$$S_{ij,t+1}=\left[\beta\frac{\gamma_i}{\gamma_j}\alpha_i^s\right]\omega_{ij}Y_{jt}. \tag{60}$$

The IO coefficient matrix $\boldsymbol{\Omega}$ of Long-Plosser model is constant because sector j's optimal saving rate of intermediate good i is proportional to sector j's output elasticity of intermediate good i—ω_{ij}, in accordance with the modified Golden rule.

In sharp contrast, in our model there are two IO tables (both are time-varying): one for the optimal saving rate of intermediate goods $\widetilde{\boldsymbol{\Omega}}_t$; and another for the optimal saving rate of capital investment goods $\widehat{\boldsymbol{\Omega}}_t$. These IO matrices are time-varying because under incomplete markets the sectoral labor supply L_{jt} and leisure externalities ψ_t from the KUWJ preferences both affect the optimal saving rate (or input-output ratio) for intermediate goods $\tilde{\omega}_{ijt}$ and that for investment goods ξ^z_{ijt}.

To dissect the issue further and compare our model with the Long-Plosser model in a more straightforward manner, we can shut down capital ($\alpha^k_j = 0$ for all j), lump-sump tax ($T_{jt} = 0$ for all j), and KUWJ preferences ($\psi_t = 1$), but keep only the feature of heterogeneous money demand with incomplete markets. Then equations (56) and (57) are simplified to

$$\boldsymbol{\gamma}_t = \left(\mathbf{E} - \beta \mathbf{Diag}\left(\boldsymbol{\alpha}^{\mathbf{s}}\right) \widetilde{\boldsymbol{\Omega}}'_t\right)^{-1} \boldsymbol{\varphi}, \tag{61}$$

where the IO coefficient matrix $\widetilde{\boldsymbol{\Omega}}_t$ has components $\widetilde{\omega}_{ijt} \equiv \omega_{ij} \cdot \mathbb{E}_t \left(L_{i,t+1}/L_{it}\right)$. This equation differs from equation (58) only because $\widetilde{\omega}_{ijt} \neq \omega_{ij}$. The reason is that in the Long-Plosser model sectoral labor L_{it} is constant for all sectors $i \in \mathbf{N}$. But in our model with incomplete markets, in contrast, sectoral labor is time varying and its growth rate L_{it+1}/L_{tt} enters the optimal saving rates of intermediate goods (which constitute the IO coefficients). Labor is time varying in our model because money demand and its distribution θ^*_t are time varying. Thus, any aggregate shocks that inevitably change the household money demand and its distribution will change not only labor supply but also the economy's marginal propensities to save in intermediate goods.

The reason that the fraction of currently produced commodity j to be saved and allocated to sector i for the next period depends not only on the input-output elasticity ω_{ij} in the production technology, but also on expected labor growth in sector i, $\mathbb{E}_t \left(L_{i,t+1}/L_{it}\right)$, is as follows. Since labor and intermediate goods are complements, a higher future labor demand in sector i relative to the present implies higher productivity of all intermediate goods used in sector i in the next period; so the expected growth rate in sector i's labor affects the choices (saving rates) of intermediate goods. Hence, the optimal input-output ratio is proportional to $\widetilde{\omega}_{ijt} \equiv \omega_{ij} \cdot \mathbb{E}_t \left(L_{i,t+1}/L_{it}\right)$.

Based on the above discussions, it becomes clear why the KUWJ preference specification amplifies monetary shocks—because it reinforces the effect of a time-varying labor supply on the IO coefficients $\widetilde{\omega}_{ijt} \equiv \omega_{ij} \mathbb{E}_t \left[\frac{\psi_{t+1}}{\psi_t} \frac{L_{i,t+1}}{L_{it}}\right]$.

2.5 Labor Wedge and Labor Dynamics

Labor Wedge. Recall that in one-sector representative-agent models without frictions, firm's marginal product of labor (MPL) equals household's marginal rate of substitution (MRS). However, this is not true in the data. The literature on business cycle accounting pioneered by Chari, Kehoe and McGrattan (2007) shows that there exists a time-varying wedge between the measured MPL (i.e.,$\partial F\left(K_t, L_t\right) / \partial L_t$) and the measured MRS (i.e., $-u_l/u_c$). They show that this labor wedge accounts for essentially all of the aggregate output fluctuations in the Great Depression and the post-war period when calibrated to a standard one-sector representative-agent RBC model. In a recent empirical study, Karabarbounis (2014) further shows that the measured labor wedge comes mainly from the gap between the real wage and the household's MRS.

Motivated by the studies of Chari, Kehoe and McGrattan (2007) and Karabarbounis (2014), if we define the labor wedge in our model as the log difference between the aggregate MPL (the real wage W_t) and the aggregate MRS ($-u_{l,t}/u_{c,t} = \psi_t C_t$ under quasi-linear preference); namely, if we define the labor wedge in our model as $\tau_t^w \equiv \ln W_t - \ln(\psi_t C_t) = \ln \frac{W_t}{\psi_t C_t}$, and denote the labor-wage factor $\Delta_t \equiv \psi_t C_t / W_t$, then $\tau_t^w \equiv -\ln \Delta_t$. The following Lemma shows that the labor wedge in our model is purely a function of the distribution of money demand across households:

Lemma 1 *The labor wedge is given by*

$$\tau_t^w \equiv -\ln \Delta_t \geq 0,$$

where

$$\Delta_t = \frac{\Gamma(\theta_t^*) D(\theta_t^*)}{\varepsilon + \theta_t^*} = \frac{\mathbb{E}\max(\varepsilon+\theta, \varepsilon+\theta_t^*) \cdot \mathbb{E}\min(\varepsilon+\theta, \varepsilon+\theta_t^*)}{\varepsilon+\theta_t^*} \leq 1. \tag{62}$$

Proof: See Appendix. ■

The labor wedge τ_t^w would vanish if (i) there is no uninsurable risk (i.e., $Var(\theta) \to 0$), or (ii) the borrowing constraints do not bind (i.e., under the Friedman rule where $\theta_t^* = \theta_{\max}$), or (iii) money is not held as a store of value (e.g., under hyper inflation where $\theta_t^* = \theta_{\min}$). In each of these cases, we have $\Delta(\theta_t^*) = 1$ and $\tau_t^w = 0$. As will be shown shortly, the labor wedge τ_t^w is countercyclical in our model, as in the data.

Note that the labor wedge has three components in it and they capture the intensive margin and the extensive margin of the aggregate money demand: (i) $H(\theta_t^*)$ is the marginal propensity to hold money (as in equation (26)); (ii) $\Gamma(\theta^*) \geq 1$ is the liquidity premium of money; and (iii) θ^* is the cutoff that captures the fraction of cash-constrained households. Hence, the labor-wedge factor $\Delta(\theta_t^*)$ pertains to the economic forces behind aggregate money demand. In other words, a wedge between the marginal rate of substitution and the marginal product of labor arises in the model purely because money demand distorts the optimal allocations of household consumption and savings, consistent with the view of Friedman (1969) and Lucas (1980) that money is an inferior asset to hold.

Hence, our model suggests that an important source of the fluctuations in the measured labor wedge observed in the data by Chari, Kehoe and McGrattan (2007) and Karabarbounis (2014) could come from movements in the distribution of household money demand, which creates a time-varying gap between the marginal rate of substitution and the real returns to labor (the wage rate). This gap vanishes if and only if money demand vanishes or is no longer distortionary. In the next section, we show that the movements in the labor wedge implied by our model are consistent with the behaviors of empirically measured labor wedge documented by Karabarbounis (2014).

Labor Dynamics. Since the distribution of money demand is characterized by the cutoff θ_t^*, we can show how the time-varying nature of the distribution of money demand dictates household labor supply and hence the IO coefficients matrices in our model—recall that $\widetilde{\omega}_{ijt} \equiv \omega_{ij} \cdot \mathbb{E}_t \left(\frac{\psi_{t+1} L_{i,t+1}}{\psi_t L_{it}} \right)$. The following proposition shows that the dynamics of sectoral labor $\mathbf{L}_t \equiv \{L_{it}\}_{i \in \mathbf{N}}$ is a second-order forward-looking auto-regressive difference equation in θ_t^*:

Proposition 6 ***(Dynamics of Sectoral Labor)*** *Denote $\widetilde{L}_{jt} \equiv L_{jt}/\alpha_j^l$. If $\delta_j = \delta$ for all $j \in \mathbf{N}$, then labor dynamics in sector $j \in \mathbf{N}$ is governed by the forward second-order difference equation:*

$$\begin{aligned} (1-g_{jt})\,\widetilde{L}_{jt} &= \frac{\Delta_t - \beta(1-\delta)\,\mathbb{E}_t \Delta_{t+1}}{\psi_t}\varphi_j + \beta \mathbb{E}_t \frac{\psi_{t+1}}{\psi_t}(1-\delta)(1-g_{j,t+1})\,\widetilde{L}_{j,t+1} \\ &\quad + \beta \mathbb{E}_t \sum_{i=1}^{N} \left(\alpha_i^s \omega_{ij} + \frac{\psi_{t+1}}{\psi_t} \delta \alpha_i^k \phi_{ij} \right) \widetilde{L}_{i,t+1} \\ &\quad - \beta^2 (1-\delta)\,\mathbb{E}_t \sum_{i=1}^{N} \frac{\psi_{t+1}}{\psi_t} \alpha_i^s \omega_{ij} \widetilde{L}_{i,t+2}. \end{aligned} \tag{63}$$

Proof: See Appendix. ∎

Corollary 1 *In the absence of KUWJ preferences (i.e., $\psi_t = \psi$) and lump-sum taxes ($g_{jt} = 0$ for all j), the above dynamic labor equation can be simplified and rewritten more compactly in vector form:*

$$\begin{aligned} \widetilde{\mathbf{L}}_t &= \frac{\Delta(\theta_t^*) - \beta(1-\delta)\,\mathbb{E}_t \Delta(\theta_{t+1}^*)}{\psi} \boldsymbol{\varphi} + \boldsymbol{\rho}_1 \mathbb{E}_t \widetilde{\mathbf{L}}_{t+1} - \boldsymbol{\rho}_2 \mathbb{E}_t \widetilde{\mathbf{L}}_{t+2} \\ &\equiv \mathbb{E}_t \mathbf{A}\left(L^{-1}\right) \Delta(\theta_t^*), \end{aligned} \tag{64}$$

where $\mathcal{L}^{-1}$ is the forward operator (i.e., $\mathcal{L}^{-1} x_t = x_{t+1}$), $\mathbf{A}\left(L^{-1}\right)$ is a rational polynomial function in the forward operator, the coefficient matrices $\{\boldsymbol{\rho}_1, \boldsymbol{\rho}_2\}$ are given by

$$\boldsymbol{\rho}_1 \equiv \beta \left[(1-\delta)\,\mathbf{E} + \delta \mathbf{Diag}\left(\boldsymbol{\alpha}^k\right) \boldsymbol{\Phi} + \mathbf{Diag}\left(\boldsymbol{\alpha}^s\right) \boldsymbol{\Omega} \right]', \tag{65}$$

$$\boldsymbol{\rho}_2 \equiv \beta^2 (1-\delta)\,\mathbf{Diag}\left(\boldsymbol{\alpha}^s\right) \boldsymbol{\Omega}', \tag{66}$$

in which $\boldsymbol{\Omega} \equiv \{\omega_{ij}\}_{ij \in \mathbf{N}}$, and $\boldsymbol{\Phi} \equiv \{\phi_{ij}\}_{ij \in \mathbf{N}}$ are IO matrices for intermediate goods and investment goods, respectively.

Equation (64) suggests that optimal labor demand is forward looking—the current labor demand is a distributed sum of future changes in the distributions of money demand, discounted and compounded by the Long-Plosser input-output coefficient matrix $\mathbf{\Omega}$ for intermediate goods and that for investment goods $\mathbf{\Phi}$. This property leads to a strong effect of news shocks about future monetary policy changes through the Baumol-Tobin redistribution channel.

Also, equation (64) suggests that labor in different sectors are inter-connected, so they can be solved only jointly in vector form. In particular, as shown in equation (63), labor dynamics in sector j is affected not only by the economy's expectation of sector j's labor in the next period, but also by the expectation of other $i \in \mathbf{N}$ sectors' labor in the next and next-next periods. The forward-looking auto-regressive coefficients matrices $\{\boldsymbol{\rho}_1, \boldsymbol{\rho}_2\}$ are determined by the IO table $\mathbf{\Omega}$ for intermediate goods and the IO table $\mathbf{\Phi}$ for capital investment goods.

Also note that if there are no intermediate goods, i.e., if $\boldsymbol{\alpha}^s = 0$, then $\boldsymbol{\rho}_{2t} = 0$ and the labor dynamics is reduced to a forward-looking AR(1) process. Alternatively, if there is no capital in the model, i.e., $\delta \to 1$, and $\alpha_i^k \to 0$, then labor dynamics is also reduce to a forward-looking AR(1) process. Thus, the accumulation of capital through investment z_{ijt} and demand for intermediate goods s_{ijt} reinforce each other in propagating monetary shocks through the labor wedge factor $\Delta(\theta_t^*)$.

It is worth emphasizing again that equation (64) also indicates that labor supply is time-varying mainly because the distribution of money demand (θ_t^*) is time-varying. Labor would be constant if the labor wedge is constant, which would be true if the cutoff θ_t^* (distribution) is constant, which would be true if the variance of the idiosyncratic wealth shocks degenerates to zero ($\theta_{\min} = \theta_{\max} = \mathbb{E}(\theta)$), or if households are never borrowing constrained ($\theta_t^* = \theta_{\min}$, such as under the Friedman-rule inflation rate $\pi = \beta - 1$), or if households opt not to hold money at all ($\theta_t^* = \theta_{\max}$, such as in the case of hyper inflation where $\theta^* = \theta_{\min}$). In all such cases, the markets are complete, the labor wedge vanishes, and the labor supply becomes constant; therefore the aggregate variables in our monetary model behave very much like their counterparts in Long and Plosser (1983). This important issue will be discussed further in the following sections.

2.6 Steady-State Analysis

We define steady state as the situation without aggregate uncertainty. Hence all aggregate real variables are constant over time in the steady state. The following proposition shows that all aggregate variables in steady state can be solved analytically in a recursive manner.

Proposition 7 *In the steady state, the system of aggregate variables can be solved recursively in the following order:*

1. *Given the steady-state growth rate of money or the inflation rate* π *and the time discounting factor* β*, the cutoff* θ^* *is determined by*

$$\Gamma(\theta^*) = (1+\pi)(1+r), \tag{67}$$

where $r = 1/\beta - 1$ *and the liquidity premium function* $\Gamma(\theta^*) \equiv \int \max(\varepsilon+\theta, \varepsilon+\theta^*)\, dF(\theta)$.

2. *Sectoral labor is determined by*

$$\mathbf{L} = \mathbf{Diag}\left(\boldsymbol{\alpha}^l \boldsymbol{\gamma}'\right) \Delta(\theta^*)/\psi, \tag{68}$$

i.e., $L_j = \alpha_j^l \gamma_j \Delta(\theta^*)/\psi$ *for* $j \in \mathbf{N}$*, where the labor-wedge factor* $\Delta(\theta^*) \equiv \frac{\Gamma(\theta^*)D(\theta^*)}{\varepsilon+\theta^*}$ *and the vector of Dommar weight* $\boldsymbol{\gamma}$ *is given by*

$$\boldsymbol{\gamma} = \left(\mathbf{E} - \beta\widehat{\boldsymbol{\Omega}}'\right)^{-1}\boldsymbol{\varphi}, \tag{69}$$

and the generalized IO matrix $\widehat{\boldsymbol{\Omega}}$ *is given by*

$$\widehat{\boldsymbol{\Omega}} \equiv \mathbf{Diag}(\boldsymbol{\alpha}^{\mathbf{s}})\,\boldsymbol{\Omega} + \frac{1}{\beta}\mathbf{Diag}\left(\frac{\boldsymbol{\delta}}{\mathbf{r}+\boldsymbol{\delta}}\right)\mathbf{Diag}\left(\boldsymbol{\alpha}^k\right)\boldsymbol{\Phi} + \frac{1}{\beta}\mathbf{Diag}(\mathbf{g}), \tag{70}$$

in which $\mathbf{g}$ *and* $\frac{\boldsymbol{\delta}}{\mathbf{r}+\boldsymbol{\delta}}$ *are matrices with row components given by* g_j *and* $\frac{\delta_j}{r+\delta_j}$*, respectively.*

3. *Sectoral output and sectoral capital stock are given respectively by*

$$\ln\mathbf{Y} = \left[\mathbf{E} - \mathbf{Diag}\left(\boldsymbol{\alpha}^k\right)\boldsymbol{\Phi} - \mathbf{Diag}(\boldsymbol{\alpha}^s)\,\boldsymbol{\Omega}\right]^{-1}\ln\mathbf{Z}^y, \tag{71}$$

$$\ln\mathbf{K} = \boldsymbol{\Phi}\ln\mathbf{Y} + \ln\mathbf{Z}^{\mathbf{k}}, \tag{72}$$

where the vector $\mathbf{Y}$ *has components* Y_j*, and the vectors* $\mathbf{Z}^y$*,* $\mathbf{Z}^k$ *and* $\mathbf{Z}^s$ *are given, respectively, by*

$$\begin{aligned}
\ln\mathbf{Z}^y &= \mathbf{Diag}\left(\boldsymbol{\alpha}^k\right)\ln\mathbf{Z}^k + \mathbf{Diag}(\boldsymbol{\alpha}^s)\ln\mathbf{Z}^s + \ln\mathbf{A} + \mathbf{Diag}\left(\boldsymbol{\alpha}^l\right)\ln\mathbf{L}, \\
\ln\mathbf{Z}^k &= \mathbf{Diag}\left(\ln(\boldsymbol{\Xi}^z)\,\boldsymbol{\Phi}'\right) + \ln\boldsymbol{\epsilon} - \ln\boldsymbol{\delta}, \\
\ln\mathbf{Z}^s &= \mathbf{Diag}\left(\ln(\boldsymbol{\Xi}^s)\,\boldsymbol{\Omega}'\right),
\end{aligned}$$

in which the matrices $\boldsymbol{\Xi}^z$ *and* $\boldsymbol{\Xi}^s$ *have components given, respectively, by*

$$\xi_{ij}^z = \frac{\alpha_i^k \delta_i \phi_{ij}}{r+\delta_i}\frac{\gamma_i}{\gamma_j}, \tag{73}$$

$$\xi_{ij}^s = \frac{\alpha_i^s \omega_{ij}}{r+1}\frac{\gamma_i}{\gamma_j}. \tag{74}$$

4. *The vector of sectoral investment* $\mathbf{I}$, *vector of sectoral consumption* $\mathbf{C}$, *matrix of intermediate inputs* $\mathbf{S}=\{S_{ij}\}_{ij\in\in\mathbf{N}}$, *matrix of investment inputs* $\mathbf{Z}=\{Z_{ij}\}_{ij\in\mathbf{N}}$ *are given, respectively, by*

$$\begin{aligned} \ln\mathbf{I} &= \ln\mathbf{K}+\ln\boldsymbol{\delta}, & (75)\\ \ln\mathbf{C} &= \ln\mathbf{Y}+\ln\boldsymbol{\varphi}-\ln\boldsymbol{\gamma}, & (76)\\ \ln\mathbf{S} &= \ln\boldsymbol{\Xi}^s+\ln\overleftrightarrow{\mathbf{Y}}, & (77)\\ \ln\mathbf{Z} &= \ln\boldsymbol{\Xi}^z+\ln\overleftrightarrow{\mathbf{Y}}, & (78)\end{aligned}$$

where the matrix $\overleftrightarrow{\mathbf{Y}}$ *has identical columns given by the output vector* $\mathbf{Y}$.

5. *The prices are given by*

$$\begin{aligned} \ln\mathbf{R}^k &= \ln\boldsymbol{\alpha}^k+\ln\mathbf{Y}-\ln\mathbf{K},\\ \ln\mathbf{W} &= \ln\boldsymbol{\alpha}^l+\ln\mathbf{Y}-\ln\mathbf{L},\\ \ln\mathbf{Q} &= \boldsymbol{\varphi}'\left(\ln\boldsymbol{\varphi}+\ln\mathbf{W}\right)\mathbf{1}_{N\times1}-\ln\mathbf{W},\\ \ln W &= \boldsymbol{\varphi}'\left(\ln\boldsymbol{\varphi}+\ln\mathbf{W}\right),\\ \ln\mathbf{R}^s &= \ln\overleftrightarrow{\boldsymbol{\alpha}}^s+\ln\overleftrightarrow{\mathbf{Y}}+\ln\boldsymbol{\Omega}-\ln\mathbf{S}\end{aligned}$$

where the matrix $\overleftrightarrow{\boldsymbol{\alpha}}^s$ *has identical columns* $\boldsymbol{\alpha}^s$ *and the matrix* $\mathbf{R}^s$ *has elements* $R^s_{ij}=\alpha^s_i\omega_{ij}\frac{Y_i}{S_{ij}}$.

Proof: See Appendix. ■

Several comments are in order for Proposition 7. First, the textbook version of the Euler equation for capital accumulation in a one-sector frictionless model is given by

$$u'\left(C_t\right)=\beta\mathbb{E}_t u'\left(C_{t+1}\right)\left[\alpha Y_{t+1}/K_{t+1}+(1-\delta)\right], \tag{79}$$

and thus the steady-state aggregate saving rate is easily obtained as $\frac{I}{Y}=\delta\frac{K}{Y}=\delta\frac{1-\beta(1-\delta)}{\beta\alpha}$, where α denotes the capital share in one-sector model. Equation (73) formalizes a generalized saving formula in the presence of intermediate IO matrix (Ω) and investment IO matrix (Φ), which jointly determines the investment-to-output ratios $\xi_{ij}=Z_{ij}/Y_j$.

Secondly, there is a gap between the steady state allocation of our model and that of the Long-Plosser model due to the existence of the labor wedge $\tau^w\left(\theta^*\right)>0$. This is the case even if there is no lump-sun tax or government expenditure ($\mathbf{g}=\mathbf{0}$) and no capital ($\boldsymbol{\alpha}^k=\mathbf{0}$) such that $\widehat{\boldsymbol{\Omega}}=\boldsymbol{\Omega}$. The magnitude of this wedge, according to equation (67), is determined in the steady state by the inflation rate π. Hence, money is the key source of distortions, and inflation has permanent effects on welfare and cross-sectoral allocation of resources.

Under the Friedman rule, i.e., $1+\pi=\beta$, we have $\Gamma(\theta^*)=1$ and $\theta^*=\theta_{\min}$ according to equation (67), and consequently, the labor wedge vanishes with $\Delta(\theta^*)=1$ and $\tau^w=0$.

Also, since θ^* is bound above by $\theta_{\max}$, there must exist a maximum of inflation $\pi_{\max}$ beyond which households opt not to hold any money. This maximum inflation rate is determined under the highest possible liquidity premium $\Gamma(\theta_{\max})=\varepsilon+\theta_{\max}=(1+\pi_{\max})/\beta$, which implies $\pi_{\max}=\beta(\varepsilon+\theta_{\max})-1$.

When the inflation rate $\pi\geq\pi_{\max}$ and as a result money is no longer held as a store of value because of the excessively low real rate of return to money, we have $c(\iota)=\sum_{j=1}^{N}Q_j c_j(\iota)=(\varepsilon+\theta(\iota))W/\psi$, so consumption is completely uninsured and unsmoothed. However, the average (or aggregate) consumption is *still* given by $C=W/\psi$, the same as that under the Friedman. In this case the labor wedge also vanishes, but the welfare is at its lowest possible (minimum) level because agents can no longer smooth consumption using money. This suggests the potential *danger* of measuring the welfare cost of inflation based on aggregate variables (Wen, 2015).

Denote variables with stars as the counterparts of our model under zero labor wedges, then we have $L_j\leq L_j^*$, $Y_j\leq Y_j^*$, $I_j\leq I_j^*$, $K_j\leq K_j^*$, $C_j\leq C_j^*$ for all $j\in\mathbf{N}$ and $\widetilde{\mathbf{C}}\leq\tilde{\mathbf{C}}^*$, where $\widetilde{\mathbf{C}}$ denotes the aggregate final goods consumption (value added): $\widetilde{\mathbf{C}}\equiv\prod_{j=1}^{N}C_j^{\varphi_j}$. Specifically, we have the following relationships:

Corollary 2

$$\begin{aligned}
\ln\mathbf{L}-\ln\mathbf{L}^* &= -\tau^w\mathbf{E}, &(80)\\
\ln\mathbf{Y}-\ln\mathbf{Y}^* &= -\tau^w\left(\mathbf{E}-\mathbf{Diag}\left(\boldsymbol{\alpha}^k\right)\boldsymbol{\Phi}-\mathbf{Diag}\left(\boldsymbol{\alpha}^s\right)\boldsymbol{\Omega}\right)^{-1}\boldsymbol{\alpha}^l, &(81)\\
\ln\mathbf{K}-\ln\mathbf{K}^* &= \boldsymbol{\Phi}\left(\ln\mathbf{Y}-\ln\mathbf{Y}^*\right), &(82)\\
\ln\mathbf{I}-\ln\mathbf{I}^* &= \boldsymbol{\Phi}\left(\ln\mathbf{Y}-\ln\mathbf{Y}^*\right), &(83)\\
\ln\mathbf{C}-\ln\mathbf{C}^* &= \ln\mathbf{Y}-\ln\mathbf{Y}^*, &(84)\\
\ln\widetilde{\mathbf{C}}-\ln\widetilde{\mathbf{C}}^* &= \boldsymbol{\varphi}'\left(\ln\mathbf{Y}-\ln\mathbf{Y}^*\right). &(85)
\end{aligned}$$

Proof: See Appendix. ■

These wedges exist because holding money creates an opportunity cost for household consumption and savings, which reduce households' incentive to work. As a result, the aggregate allocation is worsened compared to the case without money or with complete markets. The magnitude of the opportunity cost depends on the magnitude and distribution of real money balances, which in turn depend on the inflation rate.

To illustrate, we adopt the calibrations specified in the next section and obtain the following figure, which presents the non-linear relationship between the labor wedge τ^w and the inflation rate π:

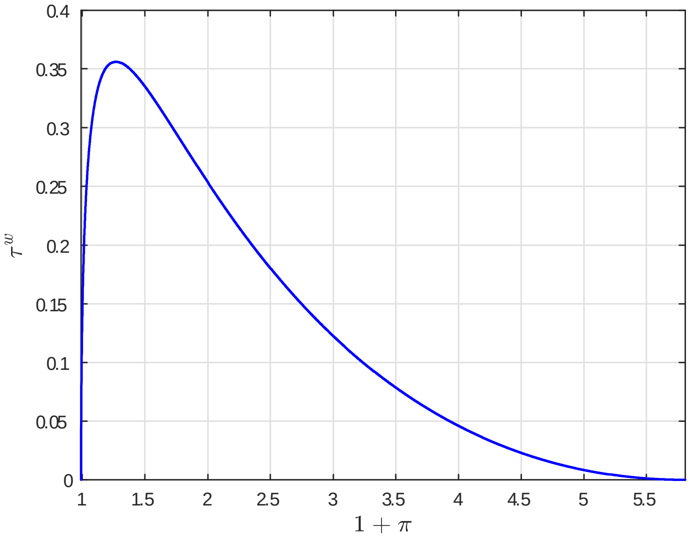

Figure 1. The Effect of Inflation π on Labor Wedge τ^w.

Clearly, the figure shows that the labor wedge vanishes either in the case of the Friedman rule with $\pi = \beta - 1$, or in the case with hyper inflation $\pi \geq \pi_{\max}$, in either cases we know that $\tau^w = 0$. The labor wedge is the highest at a moderate inflation rate around 25% per year.

The labor wedge induces wedges in other aggregate variables, such as sectoral output across all sectors, as the following figure shows. Figure 2 shows that the wedges are asymmetric across production networks—it is larger in the manufacturing sector than other sectors.

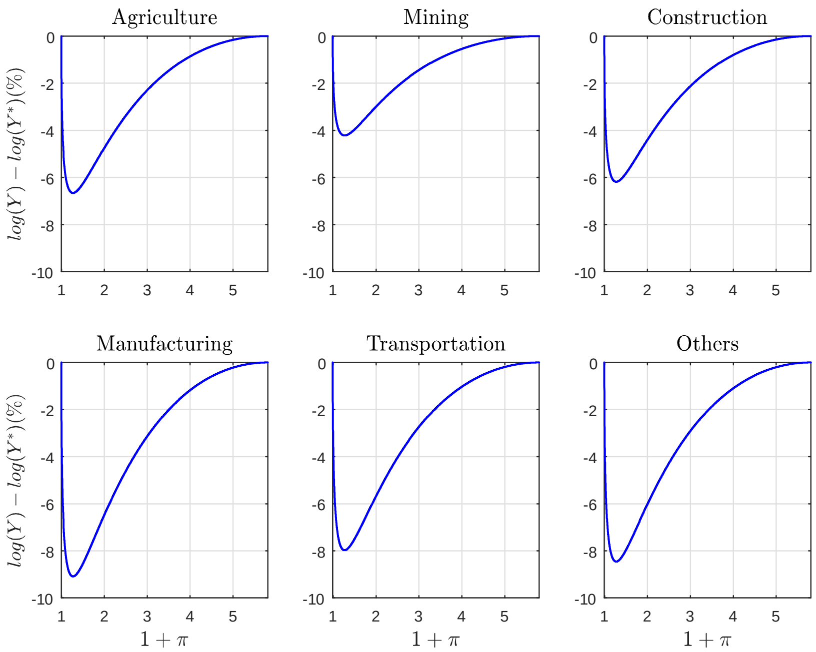

Figure 2. Heterogeneous Effects of Inflation on Sectoral Output Wedges.

3 Quantitative Analysis

3.1 Matching Distributions of Money Demand

Figure 3 plots the Lorenz curve of money demand by American households based on the Survey of Consumer Finance for the years of 1989, 1992, 1995, 1998, 2001, 2004, and 2007.[6] Each survey covers about 4000 households and is conducted every three years. The Lorenz curve shows the portion of total aggregate money balances held by different fractions of the population. The 45 degree line on the diagonal indicates complete equality. The figure shows that the amount of money held by households is highly unequally distributed, even though the top 1% of the richest households have been excluded from the sample. The implied Gini coefficients for these different years are, respectively, 0.81, 0.84, 0.83, 0.81, 0.83, 0.84, and 0.82, with an average of 0.82.[7] The degree of inequality is extremely large and has not declined over the 20 years sample period. For example, in 2007, about 22% of the population holds essentially no money (less than or equal to $10 in their checking accounts), 50% of the population holds less than 3% of the aggregate money balances, and the richest 10% of the population holds more than 75% of the total cash assets.

[6] Money demand is defined as cash and checking accounts.
[7] The average Gini coefficient without excluding the top 1% richest individuals is 0.92.

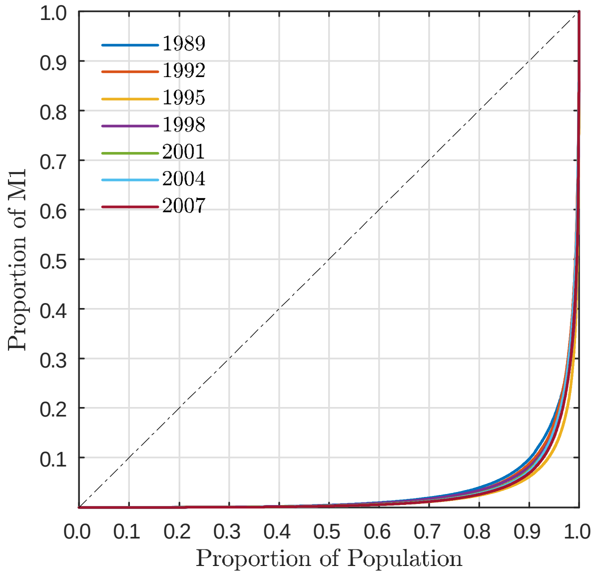

Figure 3. Distribution (Lorenz Curve) of Money Demand in the United States (1989-2007)

To capture these features of the distribution of money demand, we assume that the idiosyncratic net-worth shock $\theta(\iota)$ follows a generalized Power distribution:

$$F(\theta) = \frac{(\theta + \varepsilon)^{\sigma} - \varepsilon^{\sigma}}{(\theta_{\max} + \varepsilon)^{\sigma} - \varepsilon^{\sigma}}, \tag{86}$$

with $\sigma > 0$, and the support $\theta \in [0, \theta_{\max}]$, where $\varepsilon \in [0, 1]$ is the constant in equation (9). The mean of the distribution is $\mathbb{E}(\theta) = \frac{\sigma}{1+\sigma} \frac{(\theta_{\max}+\varepsilon)^{1+\sigma} - \varepsilon^{1+\sigma}}{(\theta_{\max}+\varepsilon)^{\sigma} - \varepsilon^{\sigma}} - \varepsilon$. Under the normalization of $\mathbb{E}(\theta) = 1 - \varepsilon$, we set $\frac{(\theta_{\max}+\varepsilon)^{1+\sigma} - \varepsilon^{1+\sigma}}{(\theta_{\max}+\varepsilon)^{\sigma} - \varepsilon^{\sigma}} = \frac{1+\sigma}{\sigma}$, which determines the upper bound $\theta_{\max}$ of the distribution. This distribution includes the Uniform distribution as a special case when $\sigma = 1$, in which case we have $\theta_{\max} = 2(1 - \varepsilon)$, and $F(\theta) = \frac{\theta}{2(1-\varepsilon)}$. In the special case where $\varepsilon = 0$, we have $F(\theta) = \left(\frac{\theta}{\theta_{\max}}\right)^{\sigma}$ and $\theta_{\max} = \frac{1+\sigma}{\sigma}$. On the other hand, when $\varepsilon = 1$, the distribution degenerates to the Dirac delta function with entire mass at $\theta = 0$, regardless of the value of σ.[8] In this case, the model reduces to a representative-agent model without idiosyncratic risks where the equilibrium value of money is $1/P = 0$, and wealth is equally distributed across agents with Gini coefficient equal to 0. Hence the distribution specified in equation (86) is quite general, and it covers a variety of interesting cases by changing the parameter values of $\{\sigma, \varepsilon\}$.

[8] The Dirac delta function is a degenerate distribution that has the value zero everywhere except at $\theta = 0$, where its value is infinitely large in such a way that its total integral is 1.

With the generalized Power distribution function, the liquidity premium and aggregate marginal propensity to consume are given by

$$\Gamma(\theta^*) = 1 + (\theta^* + \varepsilon) \frac{(\theta^* + \varepsilon)^\sigma - \varepsilon^\sigma}{(\theta_{\max} + \varepsilon)^\sigma - \varepsilon^\sigma}, \tag{87}$$
$$D(\theta^*) = 1 + (\theta^* + \varepsilon) - \Gamma(\theta^*), \tag{88}$$

then we can solve the labor wedge factor as $\Delta(\theta^*) = \frac{D(\theta^*)\Gamma(\theta^*)}{\varepsilon+\theta^*}$. Moreover, the cutoff θ^* can be solved by using equation (67), i.e., $\Gamma(\theta^*) = (1+\pi)(1+r)$. The solution is unique since Γ is monotone in θ^*.

So we can compute the model-implied Lorenz curves as follows.

Proposition 8 *At any level of wealth indexed by $\bar{\theta} \in [0, \theta_{\max}]$, the cumulative population density is $F(\bar{\theta})$; denoting the cumulative density for wealth as a fraction of aggregate wealth as $F_\omega(\bar{\theta})$, the cumulative money balances as a fraction of aggregate money balances as $F_m(\bar{\theta})$, and the cumulative consumption as a fraction of aggregate consumption as $F_c(\bar{\theta})$, then we have*

$$F_w = \frac{1}{\mathbb{E}(\varepsilon+\theta)} \int_0^{\bar{\theta}} (\varepsilon+\theta)\, dF = \frac{\sigma}{1+\sigma} \frac{(\varepsilon+\bar{\theta})^{1+\sigma} - \varepsilon^{1+\sigma}}{(\varepsilon+\theta_{\max})^\sigma - \varepsilon^\sigma} \tag{89}$$

$$\begin{aligned} F_m &= \frac{1}{H(\theta^*)} \int_0^{\bar{\theta}} \max\{\theta - \theta^*, 0\}\, dF(\theta) \\ &= \frac{1}{H(\theta^*)} \begin{cases} 0 & \text{if } \theta < \theta^* \\ \left[\frac{\sigma}{1+\sigma} \frac{(\varepsilon+\bar{\theta})^{1+\sigma} - (\varepsilon+\theta^*)^{1+\sigma}}{(\varepsilon+\theta_{\max})^\sigma - (\varepsilon)^\sigma} - (\varepsilon+\theta^*) \frac{(\varepsilon+\bar{\theta})^\sigma - (\varepsilon+\theta^*)^\sigma}{(\varepsilon+\theta_{\max})^\sigma - (\varepsilon)^\sigma}\right] & \text{if } \theta \geq \theta^* \end{cases} \end{aligned} \tag{90}$$

$$\begin{aligned} F_c &= \frac{1}{D(\theta^*)} \int_0^{\bar{\theta}} \min\{\varepsilon+\theta, \varepsilon+\theta^*\}\, dF(\theta) \\ &= \frac{1}{D(\theta^*)} \begin{cases} \left[\frac{\sigma}{1+\sigma} \frac{(\varepsilon+\bar{\theta})^{1+\sigma} - \varepsilon^{1+\sigma}}{(\varepsilon+\theta_{\max})^\sigma - \varepsilon^\sigma}\right] & \text{if } \theta < \theta^* \\ \left[\frac{\sigma}{1+\sigma} \frac{(\varepsilon+\theta^*)^{1+\sigma} - \varepsilon^{1+\sigma}}{(\varepsilon+\theta_{\max})^\sigma - \varepsilon^\sigma}\right] + \left[(\varepsilon+\theta^*) \frac{(\varepsilon+\bar{\theta})^\sigma - (\varepsilon+\theta^*)^\sigma}{(\varepsilon+\theta_{\max})^\sigma - \varepsilon^\sigma}\right] & \text{if } \theta \geq \theta^* \end{cases}. \end{aligned} \tag{91}$$

Notice that these distributions satisfy the relation, $D(\theta^*) F_c + H(\theta^*) F_m = E(\varepsilon+\theta) F_w$.

Then the Lorenz curve $\mathcal{L}$ for each of these variable is given by

$$\mathcal{L}(F_i(\theta)) = \frac{\int_0^{\bar{\theta}} dF_i(\theta)}{\int_0^{\bar{\theta}} dF(\theta)}, \text{ where } i \in \{\omega, m, c\}. \tag{92}$$

Calibration. Suppose that we set the time period to one year, $\beta = 0.96$, the annual inflation rate $\pi = 0\%$, and calibrate the parameters $\{\sigma, \varepsilon\}$ in the Power distribution function such that the implied Gini coefficient of the distribution of consumption is 0.3. These values are consistent with $\sigma = 0.01$, and $\varepsilon = 0.0015$. Under these parameter values, the implied wealth Gini is around 0.7 and money-demand Gini is around 0.8. The model-implied Lorenz curves are graphed in Figure 4.[9] These predictions are far more consistent with the US data than in the model of Dong and Wen (2019) under idiosyncratic preference shocks. Notice in Figure 4 that the distribution of money demand is far closer to that of wealth than to consumption, as is also the case in the data. This is the consequence of holding money as an asset instead of a means of payment, so money serves mainly as a buffer stock to smooth consumption against wealth (income) shocks. The better money can serve as a store of value to smooth consumption, the closer is the distribution of money demand to that of wealth than to consumption.

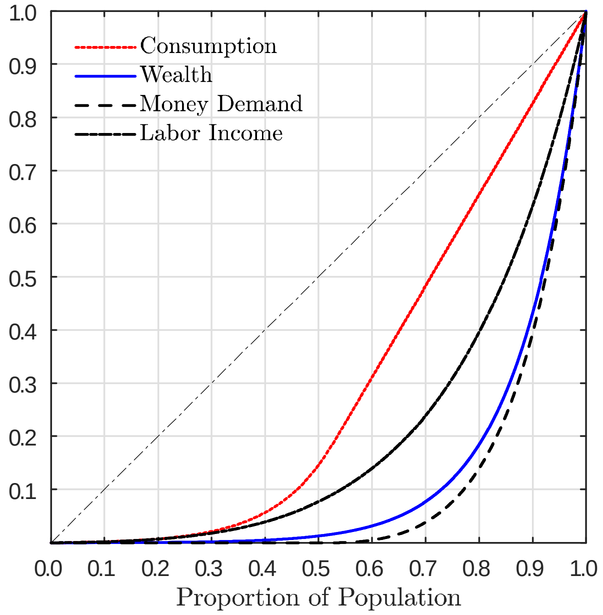

Figure 4. Predicted Lorenz Curves (Distributions) of Consumption, Money Demand, and Wealth.

Figure 4 indicates that the liquidity-constrained (cash-poor) agents are less able to smooth consumption, so the consumption curve and the wealth curve are close to each other toward the left corner of the figure, whereas the cash-rich agents are better able to smooth consumption using money, so the consumption curve and wealth curve lie far apart toward the right corner the figure. Hence, the Lorenz curve for consumption is not symmetric. In particular, the consumption level is constant across agents for the richest 45% of the population because they

[9] Since the model cannot pin down individual household level of savings for intermediate goods, we are unable to pin down individual labor income $wl(\iota)$. The Lorenz curve of labor income in Figure 4 is estimated based on a simpler model without intermediate goods; see the Appendix for details. We conjecture that it closely resembles the true Lorenz curve of labor income in the current model.

are not liquidity constrained. Therefore the consumption Lorenz curve becomes increasingly like a straight line toward the right side of the graph.

The calibrated model is also able to rationalize the empirical aggregate "money demand curve" estimated by Lucas (2000). Using historical data for GDP, money stock (M1), and the nominal interest rate, Lucas (2000) showed that the ratio of M1 to nominal GDP is downward slopping against nominal interest rate. Lucas interpreted this downward relationship as a "money demand" curve and argues that it can be rationalized by the Sidrauski (1967) model of money-in-utility. Lucas estimated that the empirical money demand curve can be best captured by a power function of the form

$$\frac{M}{PY} = \mathcal{A} \cdot r^{-\eta},$$

where $\mathcal{A}$ is a scale parameter, r the nominal interest rate, and η the interest elasticity of money demand. He showed that $\eta = 0.5$ gives the best fit. Because the money demand specified by Lucas is identical to the inverted velocity, a downward-sloping money demand curve is the same as an upward-sloping velocity curve (namely, velocity is positively related to nominal interest rate or inflation).

Similar to Lucas model, the money demand curve implied by our model takes the form

$$\frac{M}{PY} = \mathcal{A} \cdot \frac{H(\theta^*)}{D(\theta^*)},$$

where $\mathcal{A}$ is a scale parameter, the functions $\{H, D\}$ are defined in Proposition 4, and the cutoff θ^* is a function of the nominal interest rate implied by equation (67). Similar to Wen (2015), our theoretical model delivers a close fit to the US data, and therefore the figure is not reported.

3.2 Business-Cycle Analysis (I): 6 Sectors and No Capital

To start with, we compare our current wealth-shock setup with Long and Plosser (1983) and Dong and Wen (2018a), both of which considers a 6-sector model, and no capital accumulation, i.e., $\alpha_j^k = 0$ for all $j \in \mathbf{N}$. We use the US input-output (IO) table to calibrate the input-output elasticity parameters ω_{ij} in the Cobb-Douglas production function. Table 1 shows that the manufacturing sector supplies most of its output to all the other sectors as inputs, while the construction sector relies heavily on other sectors' output as its inputs. Thus, the upstream manufacturing sector has a strong supply-push effect on the economy while the downstream construction sector has a strong demand-pull effect on the economy. We follow Long and Plosser (1983) by reducing the 15×15 IO table to a smaller IO table with $N = 6$ sectors (i.e., 1. Agriculture, 2. Mining, 3. Construction, 4. Manufacturing, 5. Transportation, 6. Other). The condensed 6×6 IO table is reported in Table 1.

Table 1. Input-output coefficients from 2007 data

	Agri.	*Min.*	*Const.*	*Manuf.*	*Trans.*	*Other*
Agri.	0.2894	0.0083	0.0300	0.2823	0.1294	0.0897
Min.	0.0005	0.2548	0.0566	0.1691	0.0690	0.1826
Const.	0.0012	0.0635	0.0117	0.2903	0.1328	0.1045
Manuf.	0.0477	0.0981	0.0219	0.4340	0.0915	0.1130
Trans.	0.0004	0.0020	0.0161	0.0890	0.1124	0.3165
Other	0.0008	0.0024	0.0309	0.0792	0.0284	0.3405

We assume that all aggregate shocks follow AR(1) processes with persistence $\rho = 0.9$. We discuss below the impulse responses of the model to each aggregate shock in turn. We start with $\nu = 0$, i.e., no labor externality. We conclude this part by addressing the role played by labor externality in amplification.

1. Sectorial TFP shocks: $\ln A_{jt} = \rho A_{j,t-1} + \varepsilon_t^{A_j}$.

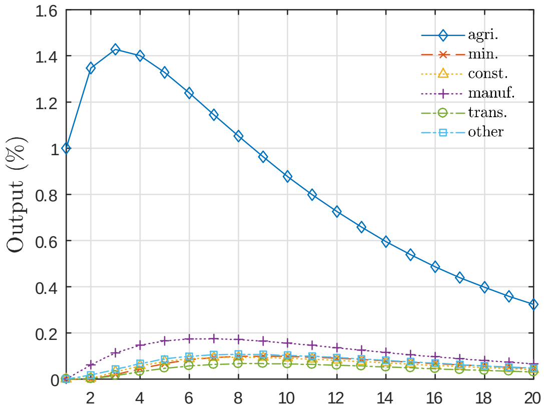

Figure 6a. TFP shock to Agriculture

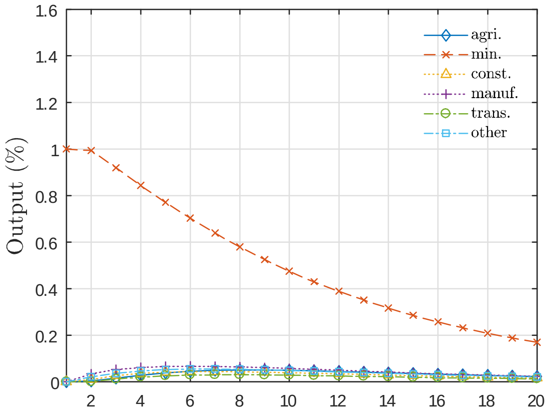

Figure 6b. TFP shock to Mining

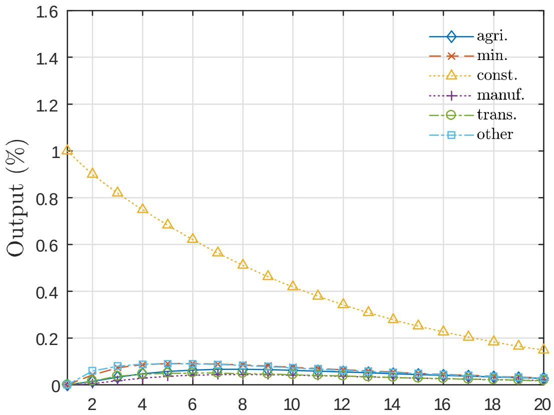

Figure 6c. TFP shock to Construction

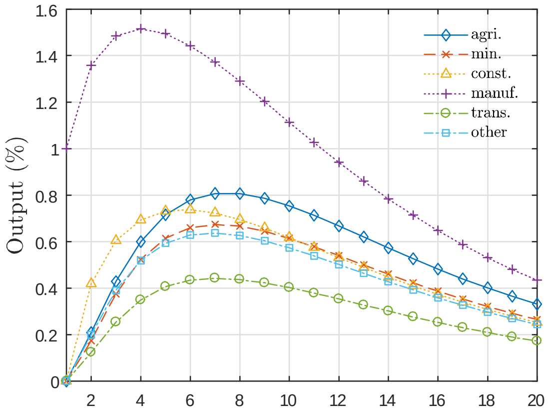

Figure 6d. TFP shock to Manufacturing

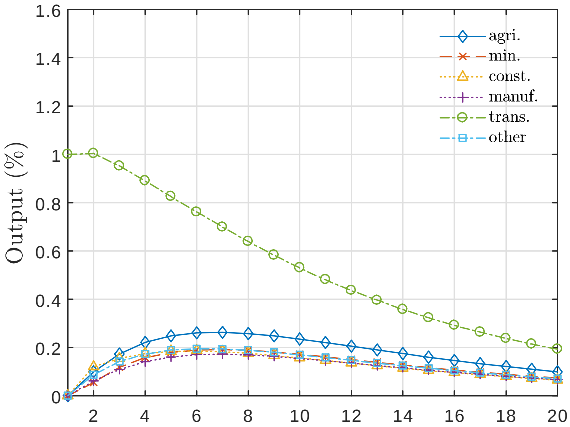

Figure 6e. TFP shock to Transp.

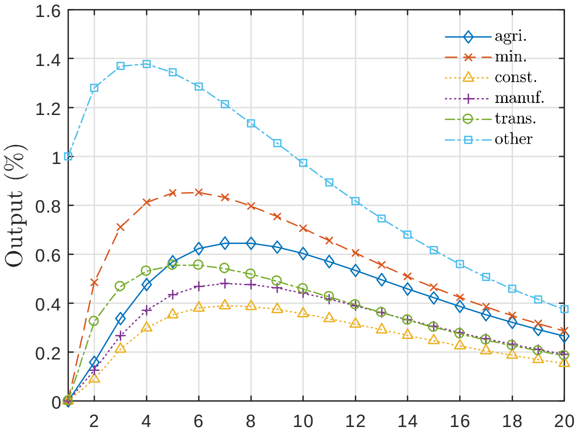

Figure 6f. TFP shock to Other Sectors

Figure 5. Sectoral TFP Shocks

Recall that labor L_{it} is constant in the Long-Plosser model. In contrast, labor is time varying in our model due to time varying distribution of money demand. Therefore there are more amplifications. The panels in Figure 5 shows that while most sectoral TFP shocks mainly impact on the sector's own output, the manufacturing sector is different (see Figure 5d): its own TFP shock exerts big influences on all of the other sectors because they all depend on the manufacturing sector's output as their inputs. Hence, sectoral TFP shocks to the manufacturing sector generates a large supply-push effect on the rest of the economy.

However, since the distribution of money demand does not respond significantly to sectoral TFP shocks in our Bewley-Lucas type model, labor is essentially constant—with a magnitude in the order of 10^{-3}, so the responses of sectorial output to TFP shocks look very similar to those in the Long-Plosser model. To gain intuition, recall that hours worked in our model differ from those in the Long-Plosser model only by the wedge $\Delta(\theta_t^*)$, which depends only on the distribution of money demand or the cutoff θ_t^*, so we plot the impulse responses of θ_t^* and the labor-wedge factor $\Delta(\theta_t^*)$, respectively, under six sectoral TFP shocks in Figure 6a and Figure 6b (top left and bottom left panels, respectively). The graphs show that both the cutoff θ_t^* and the labor-wedge factor Δ_t change very little (in the order of 10^{-15}) in response to sectoral TFP shocks. As a result, the distribution of money demand and labor supply remain approximately constant, suggesting that the income effect and substitution effect of TFP shocks on labor supply nearly cancel each other, similar to the Long-Plosser model.

This result may lead to the incorrect conclusion that heterogeneity and market incompleteness do not matter for understanding aggregate fluctuations (as argued by Krusell and Smith, 1998). In sharp contrast, the panels in the right columns in Figures 4a and 4b show that the cutoff θ^* and the labor-wedge factor $\Delta(\theta_t^*)$ increase dramatically under a monetary shock (we defer the specification of monetary shocks to the next subsection), with an order of magnitude 10^{15} times that under TFP shocks. Namely, a 10% increase in the money supply induces a 6% increase in the cutoff and a 0.16% increase in the labor-wedge factor, compared with a tiny 2.5×10^{-15} percent increase in the cutoff and a similar change in the labor wedge under TFP shocks, suggesting a significantly larger multiplier effect of demand-side shocks than supply-side shocks. More importantly, the labor-wedge factor $\Delta(\theta_t^*)$ is procyclical, suggesting that the labor wedge $\tau_t^w = -\log\Delta(\theta_t^*)$ is countercyclical, as in the data.

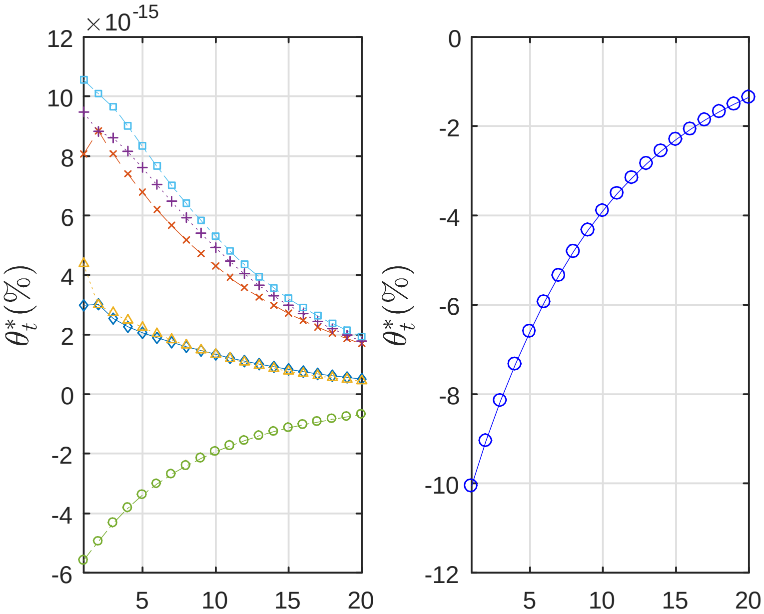

Figure 7a. Impulse response of θ_t^* under sectoral TFP shocks (left) and monetary shock (right), and the legend are referred to that in Figure 4b.

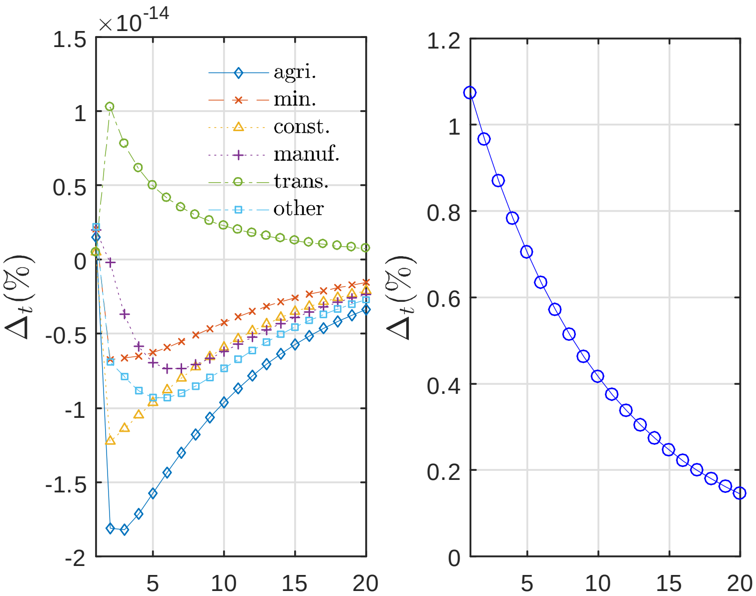

Figure 7b. Impulse response of the labor-wedge factor $\Delta(\theta_t^*)$ under sectoral TFP shocks (left) and monetary shock (right).

Since the empirically measured labor wedge is the dominant factor explaining the business cycle in the data, and since our model-implied labor wedge $\tau^w(\theta_t^*) \equiv -\log \Delta(\theta_t^*)$ is far more volatile and does a better quantitative job of matching the data-implied labor wedge under monetary shocks than under TFP shocks, our model lends support to Ramey's (2016) observation that monetary policy shocks are central to our understanding of the business cycle.

2. Monetary shock

Money is not neutral in our model. To see this, we assume that the aggregate money stock is stationary around the mean $\bar{M}$: $\bar{M}_{t+1} = \bar{M} + \bar{\tau}_t$, where money injection $\bar{\tau}_t$ follows an AR(1) process:

$$\bar{\tau}_t = \rho_\tau \bar{\tau}_{t-1} + \varepsilon_t^\tau. \tag{93}$$

Such a specification implies that any injected money is eventually taken out of the economy, as in the US qualitative easing episodes after the recent financial crisis.[10]

The left panel in Figure 7 shows the responses of the aggregate money stock M_{t+1} (red triangles) and the aggregate price level P_t (blue circles). Clearly, the aggregate price level does not respond to the money supply one-for-one: a 10% increase in the money stock causes only about a 2% increase in the price level in the impact period, as if prices are sticky despite flexible prices in our model. The sluggish response in the price level implies that the velocity of money (V_t) declines, as shown in the middle panel. A persistently declining velocity of money also suggests a persistent decrease in the liquidity premium $\Gamma(\theta_t^*)$, which captures the persistent liquidity effect of money observed in the data and helps solve a long-standing puzzle in monetary theory regarding the liquidity effect of money (see, e.g., Christiano, Eichenbaum, and Evans, 1999, for a literature review on this subject). The right panel shows that the cutoff θ_t^* increases significantly on impact and then declines slowly over time under the monetary injection, suggesting that household real money demand increases sharply and remains high and the probability of a binding liquidity constraint $\Pr[\theta \geq \theta_t^*]$ drops.

Most importantly, money has real effects on sectoral output Y_{it} for all $i \in \mathbf{N}$, as shown in Figure 8a. Notice the endogenous multiplier-accelerator effect of money supply shocks in our incomplete-market multi-sector economy: a 10% increase in the money supply can generate a non-trivial response in sectoral output across all sectors, with the typical hump-shaped pattern observed in the data. In the manufacturing sector and agricultural sector, the peak response is reached only 4 quarters after the shock. The responses of the construction and transportation sectors are the strongest, while the agriculture and manufacturing sectors are the weakest, in contrast to the case of sectoral TFP shocks.

[10]Since aggregate money demand follows the law of motion, $M_{t+1} = M_t + \tau_t$, then the money market clearing condition, $M_{t+1} = \bar{M}_{t+1}$, implies that cash received by households each period is given by $\tau_t = \bar{\tau}_t - \bar{\tau}_{t-1}$, which has an ARMA(1,1) representation: $\tau_t = \rho_\tau \tau_{t-1} + \varepsilon_t^\tau - \varepsilon_{t-1}^\tau$, suggesting that aggregate money demand is also stationary.

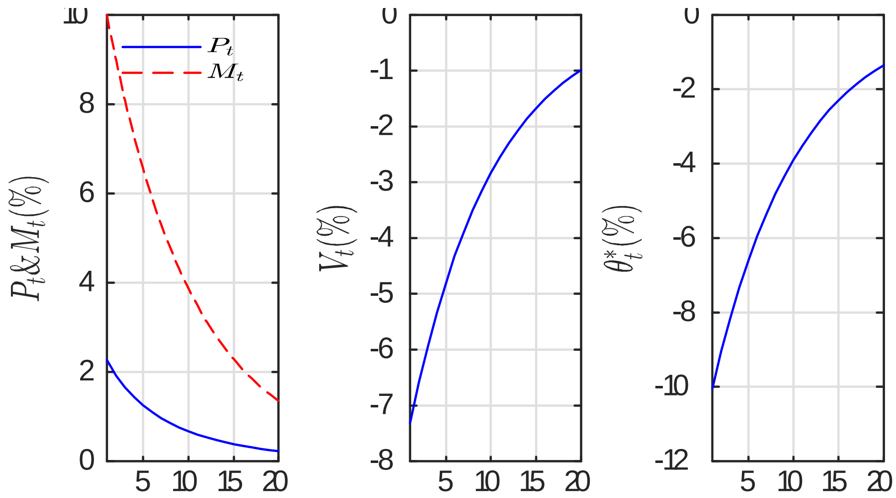

Figure 8. From left to right: impulse responses of the price level (P_t), velocity (v_t), and the cutoff (θ_t^*) to monetary shock.

The reason of such asymmetric affects across sectors are suggested by the input-output table (Table 1). The manufacturing sector supplies output to all sectors (including its own) as shown by the significantly large input-output coefficients (the column entries are relatively large), but does not require many inputs from other sectors (the row entries are relatively small); hence a TFP shock to this upstream sector has a strong "supply-push" effect on the entire economy (as noted before). On the other hand, the construction (and transportation) sector uses many other sectors' output as its own inputs (the row entries are relatively large) but is not the main provider of inputs to other sectors (the column entries are relatively small), so this downstream sector has a strong "demand-pull" effect on the entire economy. So monetary shocks act like aggregate demand shocks, enticing households to increase savings more proportionately on commodities produced by the downstream sector(s) than on commodities produced by the upstream sector(s). As a result, the responses from the upstream sector(s) (such as agriculture and manufacturing) are less volatile but more persistent over time because of delays. A similar rank of sectoral labor responses to monetary shock is revealed in Figure 8b.

Such a monetary non-neutrality originates from the distributional effect of money in the economy: only those households with a binding borrowing constraint will respond to the monetary injection by significantly increasing consumption—because of the relaxation of liquidity shortages, while liquidity-abundant households would hoard the injected money instead of spending it; thus, the aggregate price level does not respond one-for-one to the monetary increase, leading to higher aggregate real demand and output (amplified by sectoral labor demand). The hump-shaped propagation mechanism derives from the input-output linkages amplified by the time-varying nature of the input-output ratios $\mathbf{a}_{ijt}$. As a result, the downstream sectors (i.e., construction and transportation) that use other sectors' output the most as inputs will respond to the money injection more sharply than the upstream sectors (i.e., agriculture and manufacturing) that provide output as inputs, but the responses from the upstream sectors are more persistent and hump-shaped than the downstream sectors because of a dynamic priority ordering of household saving ratios on intermediate goods. Such asymmetric effects happen because of the asymmetric nature of the input-output network (rows vs. columns in the IO Table). The time-varying distribution of money demand also helps amplify the asymmetric feature of the IO table, through time-varying labor demand..[11]

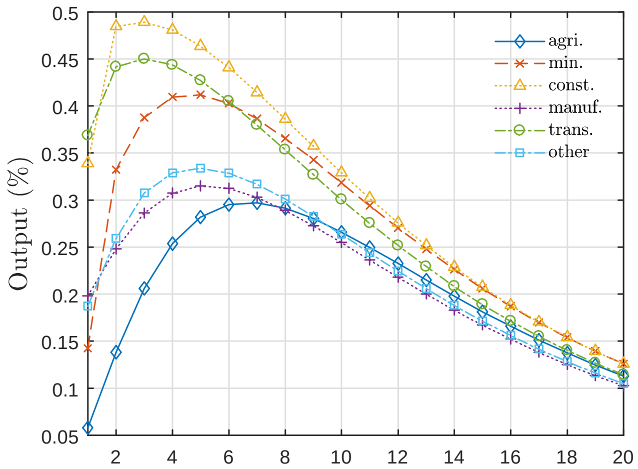

Figure 9a. Impulse response of sectoral output to a monetary shock.

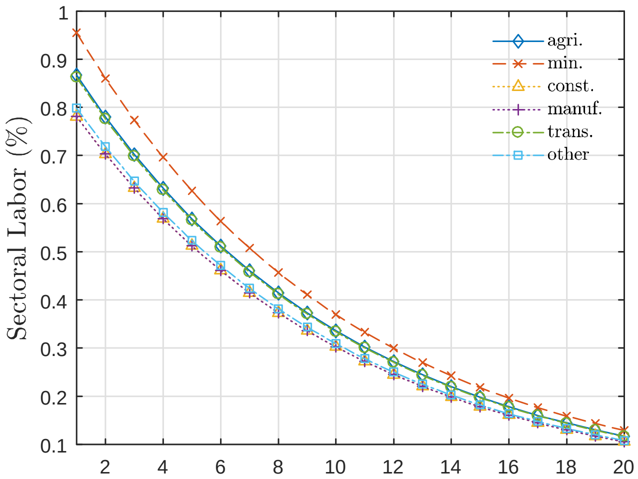

Figure 9b. Impulse responses of sectoral labor to a monetary shock.

3. Sectoral government spending shock

[11] In other words, under monetary shocks the magnitude of fluctuations is larger in a multi-sector model than in a one-sector model.

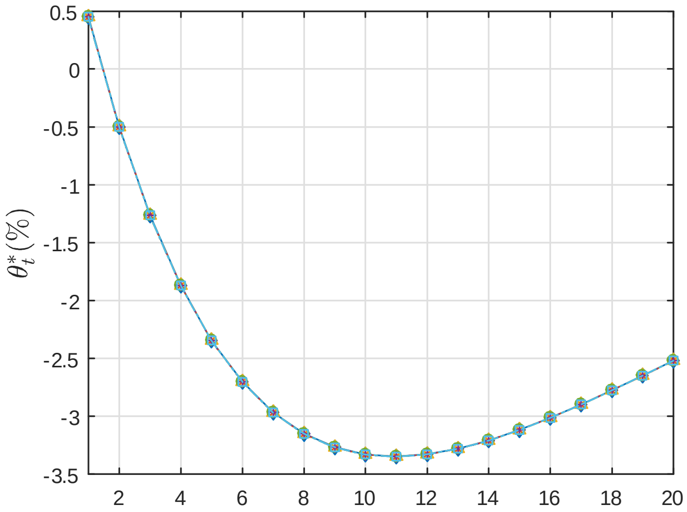

Figure 1: Figure 10. Impulse responses of the cutoff (θ_t^*) to various sectoral government spending shocks.

This subsection considers this policy question: If the government can choose the types of goods to purchase, which sector or sectors should it target to maximize the fiscal multiplier? In theory, the government could spread the budget evenly across all sectors or simply concentrate on one or a few sectors. The answer to this question obviously depends on the structure of the input-output network and is thus the subject of study here. Let the government spending shocks follow a log-linear AR(1) process:

$$\hat{G}_{jt} = \rho_g \hat{G}_{j,t-1} + \varepsilon_t^g \text{ for } i \in \mathbf{N}.$$

Figure 9 shows that the impulse responses of the cutoff to sectoral government spending shocks are identical across sectors, suggesting that sectoral government spending has the same dynamic effects on the distribution of household money demand regardless on which sector the spending is targeted.

However, a uniform change in the distribution of household money demand does not imply uniform changes in the sectoral labor and output. Equation (**??**) suggests that the input-output coefficient matrix also helps shape the dynamic responses of labor to aggregate shocks. Figure 10 shows that a sectoral government spending shock has the strongest employment effect on the targeted sector.

However, the impact of government spending on the other sectors follows the supply-push mechanism discussed above, instead of the demand-pull mechanism such that everything else equal, the upstream sector reacts more sharply to a government spending shock than the downstream sector. The intuition is as follows. Although a government spending shock is a demand-side shock, unlike a money-supply shock, the higher demand for sector i's output is "taxed" away by the government instead of being consumed or saved by households. As a result, rational households opt to dramatically increase the labor supply to the sector most affected by government spending, to minimize the adverse impact of the government spending on the rest of the economy through the sectoral linkages. In other words, households treat government spending shock as a negative income shock, in contrast to a monetary shock that has a positive income effect. Hence, the upstream sectors such as manufacturing will respond to government spending shocks more strongly than downstream sectors such as mining and construction, to mitigate the adverse impact of the shock on the entire economy.

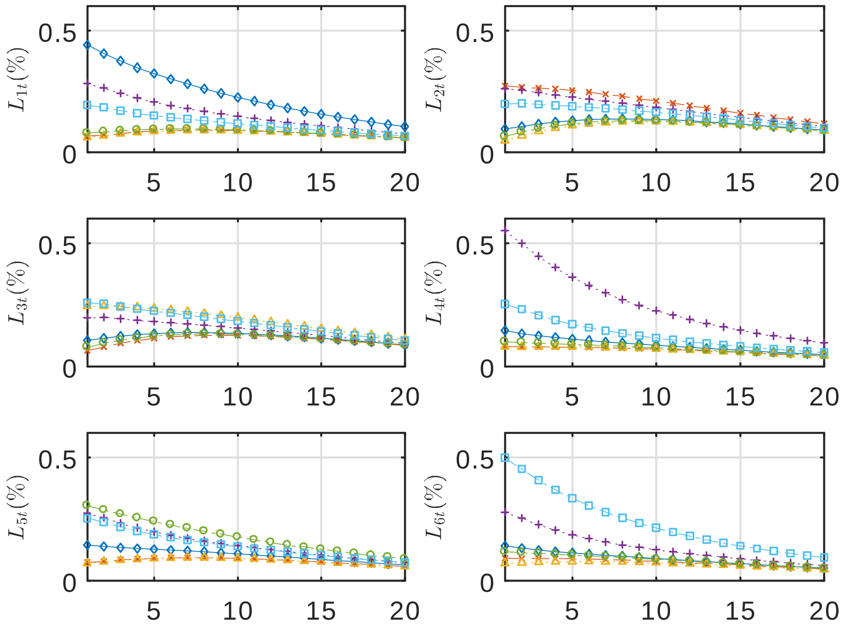

Figure 11. Impulse responses of sectoral labor (from L_{1t} to L_{6t}) to sectoral government spending shocks, where $\Diamond$ denotes Agriculture, $\times$ denotes Mining, $\triangle$ denotes Construction, $+$ denotes Manufacturing, $\circ$ denotes Transporation, and $\Box$ denotes other.

In other words, the demand-side "pulling" mechanism does not shed light on the size of the fiscal multiplier on aggregate output. Figure 11 shows the impulse responses of aggregate output to sectoral government spending shocks and it confirms this point by showing that the overall multiplier effect of sectoral government spending on aggregate output is the strongest if the government targets the upstream manufacturing sector instead of the downstream construction sector. Clearly, the effect of government spending shocks on aggregate output is the strongest when applied to the manufacturing sector and the weakest when applied to the mining and transportation sectors. These results suggest that war-time spending on military equipment may have a stronger multiplier effect (through manufacturing) than peace-time spending on

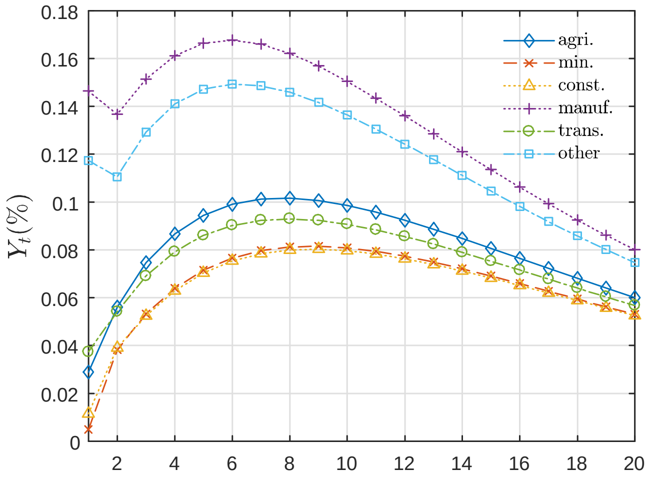

Figure 2: Figure 12. Impulse responses of aggregate output (Y_t) to sectoral government spending shocks.

infrastructure (through construction and transportation).

4. The amplification role of labor externality

We have so far set the externality parameter $\nu = 0.5$. Figure 11 shows the impulse responses of aggregate output to a 10% transitory increase in the money stock. For comparison, the solid blue line represents the case with $\nu = 0$. Cochrane (1998) uses VAR to measure the effect of money on output. He finds that (i) the output responses are protracted, hump-shaped and large; output peaks two years after the shock, and takes five years to die out. (ii) Output rises by about 5% following a 10% increase in money supply. In our model with external habit formation in leisure, the increase in aggregate output is around 1.2% at the peak when $\nu = 0.5$, as opposed to 0.5% when $\nu = 0$. However, if we set $\nu = 0.95$, then the peak response of output jumps to 12%. Therefore, the current model has the full potential to match the magnitude of monetary non-neutrality found in the data without appealing to sticky prices.

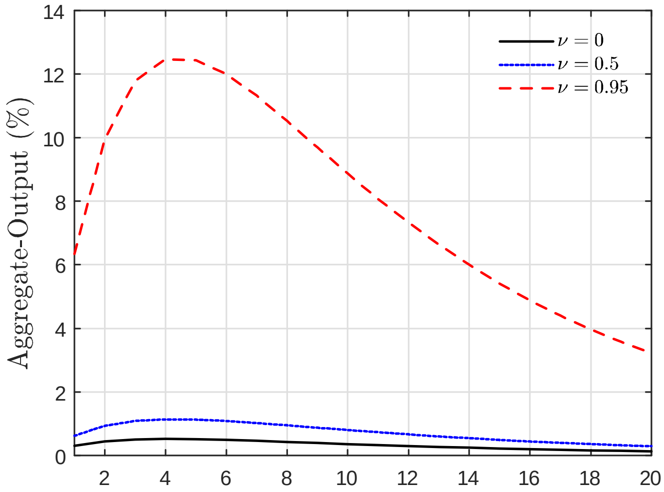

Figure 11. Effects of Monetary Policy Shock on Aggregate Ouput with Labor Externality.

3.3 Business-Cycle Analysis (II): 14 Sectors with Capital Accumulation

Table 2. Sectoral Parameters

Sector	capital share	labor share	intermediate share	final use
	(α_i^k)	(α_i^l)	(α_i^s)	(φ_i)
Agriculture	0.18	0.24	0.58	0.01
Mining	0.39	0.23	0.38	0.01
Utilities	0.41	0.17	0.42	0.02
Construction	0.13	0.39	0.48	0.07
Manufacturing	0.16	0.20	0.64	0.20
Wholesale	0.33	0.37	0.30	0.04
Retail	0.27	0.41	0.32	0.10
Transportation	0.16	0.35	0.49	0.02
Information	0.30	0.25	0.45	0.04
FIRE	0.48	0.24	0.28	0.15
PBS	0.13	0.50	0.37	0.06
Education	0.08	0.53	0.39	0.17
Arts	0.18	0.38	0.44	0.07
Other Services	0.13	0.49	0.38	0.04

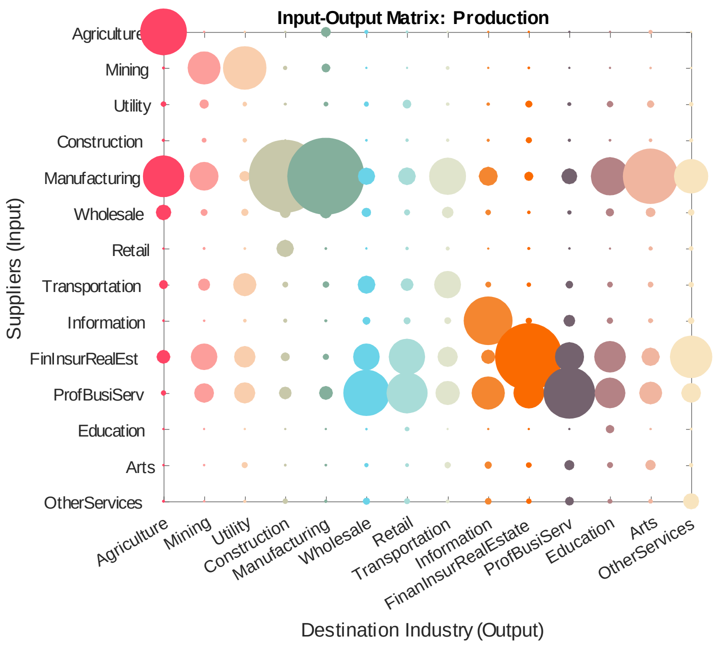

Figure x.x: IO Matrix (Normalized): a typical element ω_{ij}

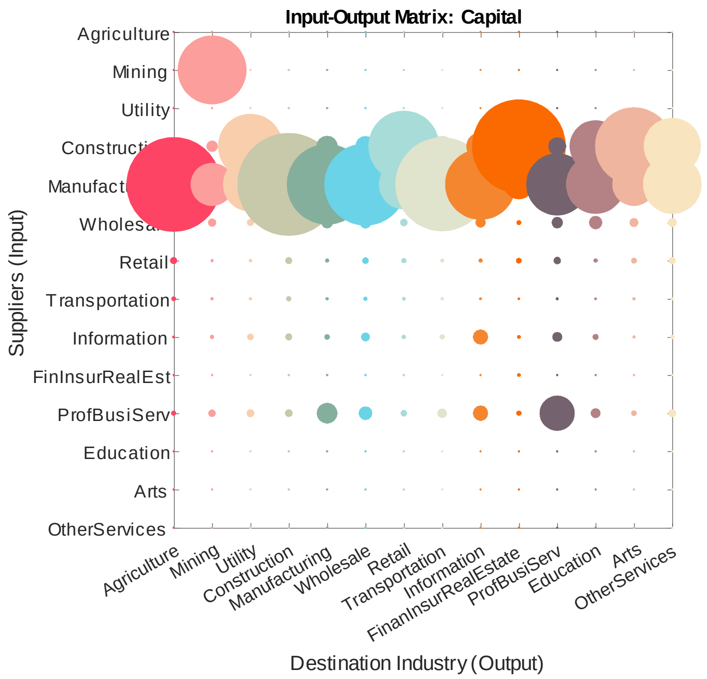

Figure x.x: IO Matrix (Normalized): a typical element ω_{ij}

We do the robust analysis by switching from 6 to 14 sectors. To start with, we normalize the sectoral TFP and sectoral investment shock such that $A_j = \epsilon_j = 1$ for all $j \in \mathbf{N}$. Since IO table is time varying, and the original parameterization using 6 sector in Long and Plosser (1983) might miss too much details of other sectors. Therefore we switch to 2-digit IO table with 15 sectors, including government. However, since government cannot be simply interpreted as profit-maximization sector, we remove government from the IO table. That is, now $N = 14$. We calibrate the labor share, capital share, and the intermediate share of the 14 sectors. See Table 2 for the summary. Moreover, in the following two figures, we illustrate the normalized IO table (with typical element ω_{ij}) and the gross IO table (with typical element $\alpha_i^s \omega_{ij}$) respectively. Table reveals that the sectors of Manufacturing, FIRE (Finance, Insurance, and Real Estate), and PBS (Professional & Business Service) are the top 3 for their outdegree. See Philippon and Reshef (2012) and Buera and Kaboski (2012) respectively for their documentation of the rise of the finance, and that of the service sector of US economy.

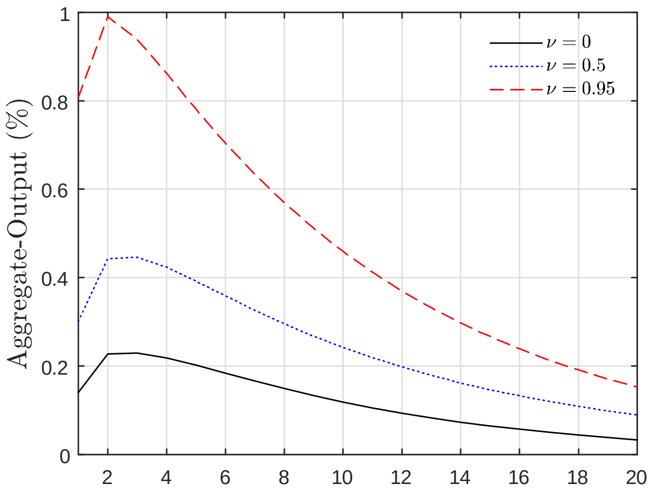

Figure x.x. Effects of Monetary Policy Shock on Aggregate Ouput with Labor Externality.

4 Conclusion

We build an analytically tractable model of dynamic production networks with incomplete insurance markets and heterogeneous money demand. We use the model to quantify the classic Baumol-Tobin redistribution channel of monetary policy. Our model can explain (i) the joint distribution of household consumption and money demand and (ii) the strong propagation mechanism of monetary shocks for the business cycle found in empirical VARs across production sectors. We show that the Baumol-Tobin redistribution channel of monetary non-neutrality can be greatly magnified and propagated through endogenous leisure choices and production networks. Our model can account for the hump-shaped impulse responses of sectoral output and employment to monetary shocks, thanks to the endogenous linkage between the distribution of household money demand and firms' input-output coefficient matrix. Our model provides an alternative framework to the Heterogeneous Agent New Keynesian (HANK) model for monetary policy analysis.

References

Acemoglu, D., Akcigit, U. and Kerr, W., 2016. Networks and the macroeconomy: An empirical exploration. *NBER Macroeconomics Annual*, 30(1), pp. 273-335.

Acemoglu, D. and Azar, P.D., 2017. Endogenous Production Networks (No. w24116). National Bureau of Economic Research.

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

Acemoglu, D., Ozdaglar, A. and Tahbaz-Salehi, A., 2017. Microeconomic origins of macroeconomic tail risks. *The American Economic Review*, 107(1), pp.54-108.

Akyol, A., 2004. Optimal monetary policy in an economy with incomplete markets and idiosyncratic risk. *Journal of Monetary Economics*, 51(6), pp.1245-1269.

Ando, S., 2014. Measuring US sectoral shocks in the world input-output network. *Economics Letters*, 125(2), pp.204-207.

Atalay, E., 2017. How important are sectoral shocks? *American Economic Journal: Macroeconomics*, 9(4), pp.254-80.

Baqaee, D.R., 2017. Cascading failures in production networks. *Econometrica*, forthcoming.

Baqaee, D.R. and Farhi, E., 2017. The Macroeconomic Impact of Microeconomic Shocks: Beyond Hulten's Theorem (No. w23145). National Bureau of Economic Research.

Bewley, T., 1980. The optimum quantity of money. *Models of monetary economies*, ed. by John Kareken and Neil Wallace. Minneapolis, Federal Reserve Bank of Minnesota, pp. 169-210

Bewley, T., 1983. A difficulty with the optimum quantity of money. *Econometrica*, 51(5), pp.1485-1504.

Bigio, S. and La'O, J., 2016. Financial frictions in production networks (No. w22212). National Bureau of Economic Research.

Chari, V.V., Kehoe, P.J. and McGrattan, E.R., 2007. Business cycle accounting. *Econometrica*, 75(3), pp.781-836.

Chien, Y. and Wen, Y., 2017, "Optimal Ramsey Capital Income Taxation —A Reappraisal," Federal Reserve Bank of St. Louis Working Paper 2017-024C.

Chiu, J., 2014. Endogenously Segmented Asset Market in an Inventory-Theoretic Model of Money Demand. *Macroeconomic Dynamics*, 18(2), pp.438-472.

Christiano, L.J., Eichenbaum, M. and Evans, C.L., 1999. Monetary policy shocks: What have we learned and to what end?. *Handbook of Macroeconomics*, 1, pp.65-148.

Deaton, Angus, 1991, Saving and liquidity constraints, *Econometrica*, 59(5), pp.1221-1248.

Dupor, B., 1999. Aggregation and irrelevance in multi-sector models. *Journal of Monetary Economics*, 43(2), pp.391-409.

Foerster, A.T., Sarte, P.D.G. and Watson, M.W., 2011. Sectoral versus aggregate shocks: A structural factor analysis of industrial production. *Journal of Political Economy*, 119(1), pp.1-38.

Friedman, M., 1969. The Optimum Quantity of Money. Macmillan.

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

Horvath, M., 1998. Cyclicality and sectoral linkages: Aggregate fluctuations from independent sectoral shocks. *Review of Economic Dynamics*, 1(4), pp.781-808.

Horvath, M., 2000. Sectoral shocks and aggregate fluctuations. *Journal of Monetary Economics*, 45(1), pp.69-106.

Hulten, C.R., 1978. Growth accounting with intermediate inputs. *The Review of Economic Studies*, 45(3), pp.511-518.

İmrohoroğlu, A., 1992. The welfare cost of inflation under imperfect insurance. Journal of Economic Dynamics and Control, 16(1), pp.79-91.

Jones, C.I., 2013. Misallocation, input-output economics, and economic growth. In *Advances in economics and econometrics: Tenth world congress* (Vol. 2), pp. 419-58. Cambridge, UK: University Press.

Kalemli-Ozcan, S., Kim, S.J., Shin, H.S., Sørensen, B.E. and Yesiltas, S., 2014. Financial shocks in production chains. In American Economic Association meetings, January.

Karabarbounis, L., 2014. The labor wedge: MRS vs. MPN. *Review of Economic Dynamics*, 17(2), pp.206-223.

Khan, A. and Thomas, J.K., 2007. Inventories and the business cycle: An equilibrium analysis of (S, s) policies. *American Economic Review*, 97(4), pp.1165-1188.

Kim, S.J. and Shin, H.S., 2012. Sustaining production chains through financial linkages. *The American Economic Review*, 102(3), pp.402-406.

Kydland, F.E. and Prescott, E.C., 1982. Time to build and aggregate fluctuations. *Econometrica*, 50(6), pp.1345-1370.

Lagos, R. and Wright, R., 2005. A unified framework for monetary theory and policy analysis. *Journal of Political Economy*, 113(3), pp.463-484.

Long Jr, J.B. and Plosser, C.I., 1983. Real business cycles. *Journal of Political Economy*, 91(1), pp.39-69.

Long Jr, J.B. and Plosser, C.I., 1987. Sectoral vs. aggregate shocks in the business cycle. *The American Economic Review*, 77(2), pp.333-336.

Lucas Jr. R.E., 1980. Equilibrium in a pure currency economy. *Economic inquiry*, 18(2), pp.203-220.

Lucas, R.E. Jr., 2000, Inflation and welfare, *Econometrica*, 68(2), pp. 247-274.

Luo, S., 2017. Propagation of financial shocks in an input-output economy with trade and financial linkages of firms. Working Paper, Virginia Tech.

Oberfield, E., 2017. A theory of input-output architecture. *Econometrica*, forthcoming.

Ozdagli, A. and Weber, M., 2017. Monetary policy through production networks: Evidence from the stock market (No. w23424). National Bureau of Economic Research.

Pasten, E., Schoenle, R. and Weber, M., 2016. The Propagation of Monetary Policy Shocks in a Heterogeneous Production Economy. Working Paper, Chicago Booth.

Ragot, X., 2014. The case for a financial approach to money demand. *Journal of Monetary Economics*, 62, pp.94-107.

Samuelson, P.A., 1958. An exact consumption-loan model of interest with or without the social contrivance of money. *Journal of Political Economy*, 66(6), pp.467-482.

Ramey, V.A., 2016. Macroeconomic shocks and their propagation. *Handbook of Macroeconomics*, 2, pp.71-162.

Taschereau-Dumouchel, M., 2017. Cascades and fluctuations in an economy with an endogenous production network. Working paper, Wharton.

Shea, J., 2002. Complementarities and comovements. *Journal of Money, Credit, and Banking*, 34(2), pp.412-433.

Su, H.L., 2017. Financial frictions, capital misallocation, and input-output linkages. Working Paper, NTU.

Wen, Y., 2010. Optimal money demand in a heterogeneous-agent cash-in-advance economy. Working Paper 2010-014, Federal Reserve Bank of St. Louis.

Wen, Y., 2011. Input and output inventory dynamics. *American Economic Journal: Macroeconomics*, 3(4), pp.181-212.

Wen, Y., 2015. Money, liquidity and welfare. *European Economic Review*, 76, pp. 1-24.

Appendix

A Notation Table

Table 2. Notation for Solving the Model

Notation	Typical element	Comment
Matrix ($N \times N$):		
$\mathbf{\Omega}$	ω_{ij}	intermediate I-O matrix
$\mathbf{\Phi}$	ϕ_{ij}	investment I-O matrix
$\mathbf{\Xi}$	ξ_{ij}	$\xi_{ij} = Z_{ij}/Y_j$.
$\mathbf{S}$	S_{ij}	intermediate input
$\widehat{\mathbf{\Omega}}_t$	$\widehat{\omega}_{ijt}$	the generalized I-O matrix in period t
$\widetilde{\mathbf{\Omega}}_t$	$\widetilde{\omega}_{ijt}$	the time-varying intermediate matrix
$\mathbf{E}$	$1_{\{i=j\}}$	identity matrix
Vectors ($N \times 1$):		
$\mathbf{1}$	1	vector of ones
$\boldsymbol{\alpha}^k$	α_j^k	capital share of sector j
$\boldsymbol{\alpha}^l$	α_j^l	labor share of sector j
$\boldsymbol{\alpha}^s$	α_j^s	intermediate share of sector j
$\boldsymbol{\varphi}$	φ_j	weight of sectoral consumption
$\mathbf{g}$	g_j	expenditure ratio of sector j
$\boldsymbol{\gamma}$	γ_j	expenditure ratio of sector j
$\mathbf{K}$	K_j	capital stock of sector j
$\mathbf{L}$	L_j	labor of sector j
$\tilde{\mathbf{L}}$	$\tilde{L}_j$	L_j/α_j^l, the modified labor of sector j
$\mathbf{Q}$	Q_j	goods price of sector j
$\mathbf{Q}$	Q_j	goods price of sector j
$\mathbf{W}$	W_j	wage rate of sector j

B Data Description

TBA. Calibration of Table x.x.

C Proof

Proof of Proposition 1: Denote μ_t, λ_{jt}, ζ_{jt} and υ_t as the multiplier of constraints (5) - (11) respectively, and and assume that $\mathbf{l}_{jt}$ adopts interior solutions. The Lagrangian is then given by

$$\begin{aligned}\mathcal{L}_t &= \mathbb{E}_0\left\{\widetilde{\mathbb{E}}_0\left\{\sum_{t=0}^{\infty}\beta^t\left[\sum_{j=1}^{N}\varphi_j\ln c_{jt}-\psi_t\sum_{j=1}^{N}l_{jt}\right]\right\}\right\} \\ &+\widetilde{\mathbb{E}}_0\sum_{t=0}^{\infty}\beta^t\mu_t\left[(\varepsilon+\theta_t)\,x_t-\sum_{j=1}^{N}q_{jt}c_{jt}\right] \\ &+\sum_{t=0}^{\infty}\sum_{j=1}^{N}\beta^t\lambda_{jt}\left(k_{jt}^{1-\sigma_j}\left(i_{jt}/\delta_j\right)^{\sigma_j}-k_{j,t+1}\right) \\ &+\sum_{t=0}^{\infty}\sum_{j=1}^{N}\beta^t\zeta_{jt}\left(\epsilon_{jt}\prod_{i=1}^{N}z_{jit}^{\phi_{ji}}-i_{jt}\right)+\sum_{t=0}^{\infty}\beta^t\upsilon_t\frac{m_{t+1}}{P_t} \end{aligned}\tag{C.1}$$

where x_t and $\widetilde{y}_{jt}$ are respectively given by equation (8) and (7).

$$x_t = \frac{m_t+\tau_t}{P_t}+\sum_{j=1}^{N}q_{jt}\widetilde{y}_{jt}-\frac{m_{t+1}}{P_t}-\sum_{j=1}^{N}\sum_{i=1}^{N}Q_{jt}s_{ij,t+1}-\sum_{j=1}^{N}\sum_{i=1}^{N}Q_{jt}z_{ijt}-T_t, \tag{C.2}$$

$$\widetilde{y}_{jt} = R_{jt}^{k}k_{jt}+\sum_{i=1}^{N}R_{jit}^{s}s_{jit}+W_{jt}l_{jt}. \tag{C.3}$$

The FOC on $\{c_t, l_t, s_{t+1}, z_t, i_t, k_{t+1}, m_{t+1}\}$ is given by

$$\frac{\varphi_j}{c_{jt}(\iota)} = Q_{jt}\mu_t(\iota), \tag{C.4}$$

$$\psi_t = W_t\widetilde{\mathbb{E}}_t\left\{(\varepsilon+\theta_t(\iota))\,\mu_t(\iota)\right\}, \tag{C.5}$$

$$Q_{jt}\widetilde{\mathbb{E}}_t(\varepsilon+\theta_t)\,\mu_t = \beta\mathbb{E}_t\widetilde{\mathbb{E}}_{t+1}(\varepsilon+\theta_{t+1})\,\mu_{t+1}Q_{i,t+1}R_{ij,t+1}^{s}, \tag{C.6}$$

$$Q_{jt}\widetilde{\mathbb{E}}_t(\varepsilon+\theta_t)\,\mu_t = \zeta_{it}\phi_{ij}\frac{i_{it}}{z_{ijt}}, \tag{C.7}$$

$$\zeta_{jt} = \lambda_{jt}\delta_j\frac{k_{j,t+1}}{i_{jt}}, \tag{C.8}$$

$$\lambda_{jt} = \beta\mathbb{E}_t\widetilde{\mathbb{E}}_{t+1}(\varepsilon+\theta_{t+1})\,\mu_{t+1}Q_{j,t+1}R_{j,t+1}^{k}+\beta\mathbb{E}_t\lambda_{j,t+1}\frac{\partial k_{j,t+2}}{\partial k_{j,t+1}}, \tag{C.9}$$

$$\frac{\mu_t}{P_t} = \beta\mathbb{E}_t\frac{(\varepsilon+\theta_{t+1})\,\mu_{t+1}}{P_{t+1}}+\frac{\upsilon_t}{P_t}, \tag{C.10}$$

where $W_t = Q_{jt}W_{jt}$, for all $j \in \mathbf{N}$, and equation (C.5), (C.6), and (C.7) reflect that decisions for labor supply $l_{jt}(\iota)$, intermediate input $s_{jit}(\iota)$ and investment input $z_{jit}(\iota)$ are made before the idiosyncratic wealth shock (and hence the value of $\mu_t(\iota)$) is realized. Equation (C.5) implies the law of one price implies, i.e., $Q_{jt}W_{jt} \equiv W_t$ for all $j \in \mathbf{N}$. Then using equation (C.5), by the law of iterated expectations and the orthogonality assumption of aggregate and idiosyncratic shocks, equations (C.6), (C.10), (C.7), and (C.9) can be rewritten, respectively, as

$$\frac{Q_{jt}\psi_t}{W_t} = \beta\mathbb{E}_t\frac{Q_{i,t+1}\psi_{t+1}}{W_{t+1}}R^s_{ij,t+1}, \tag{C.11}$$

$$\frac{\mu_t}{P_t} = \beta\mathbb{E}_t\frac{\psi_{t+1}}{P_{t+1}W_{t+1}} + \frac{\upsilon_t}{P_t}, \tag{C.12}$$

$$Q_{jt}\frac{\psi_t}{W_t} = \zeta_{it}\phi_{ij}\frac{i_{it}}{z_{ijt}}, \tag{C.13}$$

$$\lambda_{jt} = \beta\mathbb{E}_t\frac{\psi_{t+1}}{W_{t+1}}Q_{j,t+1}R^k_{j,t+1} + \beta\mathbb{E}_t\lambda_{j,t+1}\left(1-\delta_j\right)\frac{k_{j,t+2}}{k_{j,t+1}}. \tag{C.14}$$

where ψ_t/W_t pertains to the expected marginal utility of consumption in terms of labor, and we have used the fact that $\partial k_{j,t+2}/\partial k_{j,t+1} = \left(1-\delta_j\right)k_{j,t+2}/k_{j,t+1}$.

The decision rules for an individual's consumption and money demand are characterized by a cutoff strategy, taking as given the aggregate environment. Assuming interior solutions for labor supply and intermediate goods accumulation and in anticipation that the cutoff θ_t^* is independent of ι, we consider two possible cases below.

Case A: $\theta_t(\iota) \geq \theta_t^*$. In this case, the effect net wealth, $\theta_t(\iota)\, x_t(i)$, is relatively high. It is hence optimal to hold money as inventories to prevent possible liquidity constraints in the future. So $m_{t+1}(\iota) \geq 0$, $\upsilon_t(\iota) = 0$, and equation (C.12) reveals that

$$\mu_t(\iota) = \beta\mathbb{E}_t\frac{\psi_{t+1}P_t}{W_{t+1}P_{t+1}}. \tag{C.15}$$

In turn, equation (C.4) implies that consumption is given by

$$c_{jt}(\iota) = \frac{\varphi_j}{Q_{jt}}\left(\beta\mathbb{E}_t\frac{\psi_{t+1}P_t}{W_{t+1}P_{t+1}}\right)^{-1}. \tag{C.16}$$

Then the budget constraint (9) implies

$$\frac{m_{t+1}(\iota)}{P_t} = \left(\varepsilon+\theta_t(\iota)\right)x_t(\iota) - \sum_{j=1}^{N} q_{jt}c_{jt}(\iota) = \left(\varepsilon+\theta_t(\iota)\right)x_t(\iota) - \left(\beta\mathbb{E}_t\frac{\psi_{t+1}P_t}{W_{t+1}P_{t+1}}\right)^{-1}. \tag{C.17}$$

Since $m_{t+1}(\iota) \geq 0$, the above equation implies

$$\left(\varepsilon+\theta_t(\iota)\right) \geq \frac{1}{x_t(\iota)}\left(\beta\mathbb{E}_t\frac{\psi_{t+1}P_t}{W_{t+1}P_{t+1}}\right)^{-1} \equiv \varepsilon+\theta_t^*, \tag{C.18}$$

which defines the cutoff θ_t^*.

Case B: $\theta_t(\iota) < \theta_t^*$. In this case, the effective net wealth, $\theta_t(\iota)\, x_t(\iota)$, is relatively low. Then it is intuitively optimal for the household to spend all money in hand to smooth consumption. Therefore $\upsilon_t(\iota) > 0$, $m_{t+1}(\iota) = 0$, and we will later verify our conjecture that $\upsilon_t(\iota) > 0$. Then combining equation (C.4) and the budget constraint (9) immediately implies

$$\mu_t(\iota) = \frac{1}{\left(\varepsilon+\theta_t(\iota)\right)x_t(\iota)}, \tag{C.19}$$

and thus

$$c_{jt}(\iota) = \frac{\varphi_j}{Q_{jt}}\left(\varepsilon+\theta_t(\iota)\right)x_t(\iota). \tag{C.20}$$

Using the definition of the cutoff θ_t^* in equation (C.18), we can rewrite $c_{jt}(\iota)$ as

$$c_{jt}(\iota) = \frac{\varphi_j}{q_{jt}} \frac{\varepsilon + \theta_t(\iota)}{\varepsilon + \theta_t^*} \left(\beta \mathbb{E}_t \frac{\psi_{t+1}}{W_{t+1} P_{t+1}} \right)^{-1}. \tag{C.21}$$

Meanwhile, we can rewrite $\mu_t(\iota)$ as

$$\mu_t(\iota) = \frac{\varepsilon + \theta_t^*}{\varepsilon + \theta_t(\iota)} \left(\beta \mathbb{E}_t \frac{\psi_{t+1}}{W_{t+1} P_{t+1}} \right). \tag{C.22}$$

Substituting it into equation (C.12) yields

$$\frac{\upsilon_t(\iota)}{P_t} = \frac{\theta_t^* - \theta_t(\iota)}{\varepsilon + \theta_t(\iota)} \left(\beta \mathbb{E}_t \frac{\psi_{t+1}}{W_{t+1} P_{t+1}} \right). \tag{C.23}$$

Since $\theta_t(\iota) < \theta_t^*$, provided $P_t < \infty$, we know that $\upsilon_t(\iota) > 0$, which conforms our conjecture . Also notice that, by comparing equation (C.15) and (C.22), we know that the shadow value $\mu_t(\iota)$ is higher under case B than under case A. This is because of a tither budget constraint under case B.

In sum, unifying equation (C.15) and (C.22) suggests that

$$\mu_t(\iota) = \max\left(\frac{\varepsilon + \theta_t^*}{\varepsilon + \theta_t(\iota)}, 1 \right) \left(\beta \mathbb{E}_t \frac{\psi_{t+1}}{W_{t+1} P_{t+1}} \right). \tag{C.24}$$

In turn, equation (C.5) can be further simplified as

$$\frac{\psi_t}{W_t} = \beta \mathbb{E}_t \frac{\psi_{t+1} P_t}{W_{t+1} P_{t+1}} \Gamma(\theta_t^*), \tag{C.25}$$

where $\Gamma(\theta_t^*)$ measures the liquidity premium, or the shadow rate of return to holding money, and it is given by

$$\Gamma(\theta_t^*) \equiv \widetilde{\mathbb{E}}_t \{\max(\varepsilon + \theta_t^*, \varepsilon + \theta_t(\iota))\}. \tag{C.26}$$

Since $\mathbb{E}(\theta_t(\iota)) = 1 - \varepsilon$, we immediately know that $\Gamma(\theta_t^*) > 1$.

Moreover, the definition of the cutoff θ_t^* implies that $x_t(\iota)$ is independent of individual history. Then

$$(\varepsilon + \theta_t^*) x_t = \left(\beta \mathbb{E}_t \frac{\psi_{t+1} P_t}{W_{t+1} P_{t+1}} \right)^{-1} = \frac{W_t \Gamma(\theta_t^*)}{\psi_t}. \tag{C.27}$$

In turn, individual consumption in equation (C.4) can be rewritten as

$$c_{jt}(\iota) = \frac{\varphi_j}{Q_{jt}} \min\left\{ 1, \frac{\varepsilon + \theta_t(\iota)}{\varepsilon + \theta_t^*} \right\} (\varepsilon + \theta_t^*) x_t, \tag{C.28}$$

Meanwhile, the individual money demand is obtained as

$$\frac{m_{t+1}(\iota)}{P_t} = (\varepsilon + \theta_t(\iota)) x_t - \sum_{j=1}^{N} Q_{jt} c_{jt}(\iota) = \max\left\{ \frac{\theta_t(\iota) - \theta_t^*}{\varepsilon + \theta_t^*}, 0 \right\} (\varepsilon + \theta_t^*) x_t. \tag{C.29}$$

Proof of Proposition 4: First, we have proved previously that

$$W_t = Q_{jt} W_{jt}, \text{ for } j \in \mathbf{N}. \tag{C.30}$$

Second, integrating the policy function equation (C.28) in Proposition 1 yields

$$C_{jt} = \frac{\varphi_j}{Q_{jt}} D\left(\theta_t^*\right) X_t, \tag{C.31}$$

where $D\left(\theta_t^*\right) \equiv \mathbb{E}\left[\min\left(\varepsilon + \theta, \varepsilon + \theta^*\right)\right]$. Meanwhile, equation (C.29) implies that the aggregate money demand is given by

$$\frac{M_{t+1}}{P_t} = H\left(\theta_t^*\right) X_t, \tag{C.32}$$

where $H\left(\theta_t^*\right) \equiv 1 - D\left(\theta_t^*\right)$.

Besides, equation (C.27) reveals that $\left(\varepsilon + \theta_t^*\right) X_t = W_t \Gamma\left(\theta_t^*\right) / \psi_t$. Define $\Lambda_t = \psi_t / W_t$. Then we can rewrite the above equation as

$$\left(\varepsilon + \theta_t^*\right) X_t = \frac{\Gamma\left(\theta_t^*\right)}{\Lambda_t}. \tag{C.33}$$

Meanwhile, we have proved in Proposition 1 that

$$\frac{\psi_t}{W_t} = \beta \mathbb{E}_t \frac{P_t}{P_{t+1}} \frac{\psi_{t+1}}{W_{t+1}} \Gamma\left(\theta_t^*\right), \tag{C.34}$$

$$Q_{jt} \frac{\psi_t}{W_t} = \beta \mathbb{E}_t Q_{i,t+1} \frac{\psi_{t+1}}{W_{t+1}} R_{ijt,t+1}^s, \tag{C.35}$$

which can be rewritten as

$$1 = \beta \mathbb{E}_t \frac{\Lambda_{t+1}}{\Lambda_t} \frac{P_t}{P_{t+1}} \Gamma\left(\theta_t^*\right), \tag{C.36}$$

$$Q_{jt} = \beta \mathbb{E}_t \frac{\Lambda_{t+1}}{\Lambda_t} Q_{i,t+1} R_{ijt,t+1}^s. \tag{C.37}$$

Moreover, the production function is given by

$$Y_{it} = A_{it} K_{it}^{\alpha_i^k} L_{it}^{\alpha_i^l} \left(\prod_{j=1}^{N} S_{ijt}^{\omega_{ij}}\right)^{\alpha_i^s}, \tag{C.38}$$

The FOCs for $(L_{it}, K_{it}, S_{ijt})$ is given by

$$R_{it}^k = \alpha_i^k \frac{Y_{it}}{K_{it}}, \tag{C.39}$$

$$W_{it} = \alpha_i^l \frac{Y_{it}}{L_{it}}, \tag{C.40}$$

$$R_{ijt}^s = \alpha_i^s \omega_{ij} \frac{Y_{it}}{S_{ijt}}. \tag{C.41}$$

Household's FOC on labor supply, on the other hand, yields

$$W_t = Q_{jt} W_{jt}. \tag{C.42}$$

Moreover, equation (C.8) implies that $\frac{k_{j,t+1}}{i_{jt}}$ is the same across firms, and thus the individual law of motion of capital implies that $\frac{k_{jt}}{i_{jt}}$ is the same across firms as well. Therefore

$$k_{j,t+1}(\iota) = k_{jt}(\iota)\left(\frac{i_{jt}(\iota)}{k_{jt}(\iota)\,\delta_j}\right)^{\delta_j} = k_{jt}(\iota)\left(\frac{I_{jt}}{K_{jt}\delta_j}\right)^{\delta_j}, \tag{C.43}$$

and thus

$$K_{j,t+1} \equiv \int_0^1 k_{j,t+1}(\iota)\,d\iota = \int_0^1 k_{jt}(\iota)\left(\frac{I_{jt}}{K_{jt}\delta_j}\right)^{\sigma_j} d\iota = K_{jt}\left(\frac{I_{jt}}{K_{jt}\delta_j}\right)^{\sigma_j} = K_{jt}^{1-\sigma_j}\left(\frac{I_{jt}}{\delta_j}\right)^{\sigma_j}. \tag{C.44}$$

Similarly, equation (C.13) implies that $\frac{i_{it}}{z_{ijt}}$ is the same across firms. Therefore the individual investment can be rewritten as

$$i_t(\iota) = \epsilon_{it}\prod_{j=1}^{N} z_{ijt}^{\phi_{ij}}(\iota) = i_t(\iota)\,\epsilon_{it}\prod_{j=1}^{N}\left(\frac{z_{ijt}(\iota)}{i_t(\iota)}\right)^{\phi_{ij}} = i_t(\iota)\,\epsilon_{it}\prod_{j=1}^{N}\left(\frac{Z_{ijt}}{I_t}\right)^{\phi_{ij}} = \frac{i_t(\iota)}{I_t}\epsilon_{it}\prod_{j=1}^{N} Z_{ijt}^{\phi_{ij}}. \tag{C.45}$$

In turn, the aggregate investment is given by

$$I_{it} \equiv \int_0^1 i_{it}(\iota)\,d\iota = \int_0^1 \frac{i_t(\iota)}{I_t}\epsilon_{it}\left(\prod_{j=1}^{N} Z_{ijt}^{\phi_{ij}}\right) d\iota = \epsilon_{it}\prod_{j=1}^{N} Z_{ijt}^{\phi_{ij}}. \tag{C.46}$$

Additionally, the resource constraint in sector j is given by

$$C_{jt} + \sum_{i=1}^{N} S_{ij,t+1} + \sum_{i=1}^{N} Z_{ijt} + T_{jt} = Y_{jt}. \tag{C.47}$$

Finally, integrating (8) yields

$$X_t = \frac{M_t + \tau_t}{P_t} + \sum_{j=1}^{N} Q_{jt}\left(R_{jt}^{k}K_{jt} + W_{jt}L_{jt} + \sum_{i=1}^{N}\left(R_{jit}^{s}S_{jit} - S_{ij,t+1} - Z_{ijt}\right)\right) - T_t. \tag{C.48}$$

and the clearing condition in money market is given by

$$\bar{M}_{t+1} = M_{t+1} = M_t + \tau_t. \tag{C.49}$$

Proof of Proposition 5: Motivated by Long and Plosser (1983), we conjecture that there exists a vector of Dommar weight $\boldsymbol{\gamma}_t = \left\{\gamma_{jt}\right\}_{j\in\mathbf{N}}$ such that the policy function on consumption can be formulated as

$$\frac{C_{jt}}{Y_{jt}} = \xi_{jt}^{c} \equiv \frac{\varphi_j}{\gamma_{jt}}. \tag{C.50}$$

Substituting equation (25) into (C.50) yields the expenditure ratio of sector j as

$$\gamma_{jt} = \frac{Q_{jt}Y_{jt}}{C_t}. \tag{C.51}$$

Note that equation (C.51) can be rewritten as

$$Q_{jt} = \frac{\gamma_{jt}C_t}{Y_{jt}}. \tag{C.52}$$

Substituting equation (C.52) into (C.11) yields

$$S_{ij,t+1} = \beta\alpha_i^s\omega_{ij}\mathbb{E}_t\frac{\psi_{t+1}/\psi_t}{W_{t+1}/W_t}\frac{Q_{i,t+1}}{Q_{jt}}Y_{i,t+1} = \left(\mathbb{E}_t\frac{\psi_{t+1}/\psi_t}{W_{t+1}/W_t}\frac{\gamma_{i,t+1}C_{t+1}}{\gamma_{i,t}C_t}\right)\beta\alpha_i^s\omega_{ij}\frac{\gamma_{it}}{\gamma_{jt}}Y_{jt}. \quad (C.53)$$

Furthermore, we know that

$$\frac{\psi_{t+1}/\psi_t}{W_{t+1}/W_t}\frac{\gamma_{i,t+1}C_{t+1}}{\gamma_{i,t}C_t} = \frac{\psi_{t+1}}{\psi_t}\frac{Q_{i,t+1}Y_{i,t+1}/W_{t+1}}{Q_{it}Y_{it}/W_t} = \frac{\psi_{t+1}}{\psi_t}\frac{Y_{i,t+1}/W_{i,t+1}}{Y_{it}/W_{it}} = \frac{\psi_{t+1}L_{i,t+1}}{\psi_t L_{it}}. \quad (C.54)$$

Then by denoting $\widetilde{\omega}_{ijt} \equiv \omega_{ij}\cdot\mathbb{E}_t\left(\frac{\psi_{t+1}L_{i,t+1}}{\psi_t L_{it}}\right)$, we have

$$\frac{S_{ij,t+1}}{Y_{jt}} = \xi_{ijt}^s \equiv \beta\alpha_i^s\widetilde{\omega}_{ijt}\frac{\gamma_{it}}{\gamma_{jt}}. \quad (C.55)$$

Combining equation (C.8), (C.13), and (C.14) yields

$$Q_{it}\frac{\psi_t}{W_t}z_{jit} = \beta\delta_j\mathbb{E}_t\frac{\psi_{t+1}}{W_{t+1}}Q_{j,t+1}\phi_{ji}\alpha_j^k\frac{Y_{j,t+1}}{K_{j,t+1}}k_{j,t+1} + \beta\left(1-\delta_j\right)\mathbb{E}_t Q_{i,t+1}\frac{\psi_{t+1}}{W_{t+1}}z_{ji,t+1}. \quad (C.56)$$

Integrating both sides yields

$$Q_{it}\frac{\psi_t}{W_t}Z_{jit} = \beta\mathbb{E}_t\frac{\psi_{t+1}}{W_{t+1}}Q_{j,t+1}\delta_j\phi_{ji}\alpha_j^k Y_{j,t+1} + \beta\mathbb{E}_t Q_{i,t+1}\frac{\psi_{t+1}}{W_{t+1}}\left(1-\delta_j\right)Z_{ji,t+1}, \quad (C.57)$$

which can be rewritten as

$$\frac{Z_{jit}}{Y_{it}} = \beta\mathbb{E}_t\left\{\frac{\psi_{t+1}/\psi_t}{W_{t+1}/W_t}\frac{C_{t+1}}{C_t}\left[\frac{\gamma_{j,t+1}}{\gamma_{jt}}\frac{\gamma_{jt}}{\gamma_{it}}\delta_j\phi_{ji}\alpha_j^k + \frac{\gamma_{i,t+1}}{\gamma_{it}}\left(1-\delta_j\right)\frac{Z_{ji,t+1}}{Y_{i,t+1}}\right]\right\}, \quad (C.58)$$

where we have (C.52).

Since we have proved that $\frac{W_t}{W_{t+1}}\frac{\gamma_{i,t+1}C_{t+1}}{\gamma_{i,t}C_t} = \frac{L_{i,t+1}}{L_{it}}$, then

$$\xi_{jit}^z = \beta\mathbb{E}_t\left\{\frac{\psi_{t+1}}{\psi_t}\left[\frac{L_{j,t+1}}{L_{jt}}\frac{\gamma_{jt}}{\gamma_{it}}\delta_j\phi_{ji}\alpha_j^k + \frac{L_{i,t+1}}{L_{it}}\left(1-\delta_j\right)\xi_{ji,t+1}^z\right]\right\}, \quad (C.59)$$

where $\xi_{jit}^z \equiv Z_{jit}/Y_{it}$. Swapping the index i and j, and multiplying both sides by γ_{jt} yields

$$\xi_{ijt}^z\gamma_{jt} = \beta\mathbb{E}_t\left\{\frac{\psi_{t+1}}{\psi_t}\left[\frac{L_{i,t+1}}{L_{it}}\delta_i\phi_{ij}\alpha_i^k\gamma_{it} + \frac{L_{j,t+1}}{L_{jt}}\left(1-\delta_i\right)\xi_{ij,t+1}^z\gamma_{jt}\right]\right\}, \quad (C.60)$$

and thus Z_{ijt} can then be obtained as

$$Z_{ijt} = \xi_{ijt}^z Y_{jt}. \quad (C.61)$$

The aggregate resource constraint is given by

$$C_{jt} + \sum_{i=1}^N S_{ijt} + \sum_{i=1}^N Z_{ijt} + T_{jt} = Y_{jt}, \quad (C.62)$$

where $T_{jt} = G_{jt} = g_{jt}Y_{jt}$, where G_{jt} denotes government expenditure in sector j.

Substituting equations (C.50), (C.55), and (C.61) into (C.62) yields

$$\left(1-g_{jt}-\sum_{i=1}^{N}\xi_{ijt}\right)\gamma_{jt}=\varphi_j+\beta\sum_{i=1}^{N}\gamma_{it}\alpha_i^s\omega_{ij}\mathbb{E}_t\left(\frac{\psi_{t+1}L_{i,t+1}}{\psi_t L_{it}}\right). \tag{C.63}$$

Consequently, equation (C.60) and (C.63) jointly characterize the dynamics of $\left\{\gamma_{jt},\xi_{ijt}^z\right\}$ for $i,j\in\mathbf{N}$.

Finally, it is worth noting that equation In sum, equation (C.63) can be written in more compact way as

$$\left(\mathbf{E}-\mathbf{Diag}\left(\mathbf{g}_t+\widetilde{\boldsymbol{\xi}}_t\right)\right)\boldsymbol{\gamma}_t=\boldsymbol{\varphi}+\beta\mathbf{Diag}\left(\boldsymbol{\alpha}^{\mathbf{s}}\right)\widetilde{\boldsymbol{\Omega}}_t, \tag{C.64}$$

where a typical element of $\widetilde{\boldsymbol{\Omega}}_t$ is $\widetilde{\omega}_{ijt}\equiv\omega_{ij}\mathbb{E}_t\left(\frac{\psi_{t+1}L_{i,t+1}}{\psi_t L_{it}}\right)$, and a typical element of $\widetilde{\boldsymbol{\xi}}_t$ is $\widetilde{\xi}_{jt}\equiv\sum_{i=1}^{N}\xi_{ijt}^z$. In turn, the above equation immediately implies that

$$\boldsymbol{\gamma}_t=\left(\mathbf{E}-\beta\widehat{\boldsymbol{\Omega}}_t'\right)^{-1}\boldsymbol{\varphi}. \tag{C.65}$$

where the generalized I-O matrix is given by,

$$\widehat{\boldsymbol{\Omega}}_t\equiv\frac{1}{\beta}\mathbf{Diag}\left(\mathbf{g}_t+\widetilde{\boldsymbol{\xi}}_t\right)+\mathbf{Diag}\left(\boldsymbol{\alpha}^{\mathbf{s}}\right)\widetilde{\boldsymbol{\Omega}}_t. \tag{C.66}$$

Proof of Lemma 1: By definition, the labor wedge is given by

$$\tau_t^w\equiv\ln W_t-\ln\left(\psi_t C_t\right)=\ln\frac{W_t}{\psi_t C_t}.$$

Combining equation (30) and (47) yields

$$\frac{\psi_t C_t}{W_t}=\frac{\Gamma\left(\theta_t^*\right)D\left(\theta_t^*\right)}{\varepsilon+\theta_t^*}\equiv\Delta\left(\theta_t^*\right). \tag{C.67}$$

Therefore

$$\tau_t^w=-\ln\Delta\left(\theta_t^*\right). \tag{C.68}$$

It remains for us to prove $\tau_t^w\geq 0$, which is equivalent to proving that, for $\Delta\left(\theta^*\right)\leq 1$ for $\theta^*\in\left[\theta_{\min},\theta_{\max}\right]$. To recap, $\Gamma\left(\theta^*\right)=\mathbb{E}\left[\max\left(\varepsilon+\theta,\varepsilon+\theta^*\right)\right]$ and $D\left(\theta^*\right)=\mathbb{E}\left[\min\left(\varepsilon+\theta,\varepsilon+\theta^*\right)\right]$. Without loss of generality, we can set $\varepsilon=0$, or, equivalently, we can redefine $\theta^*=\varepsilon+\theta^*$, and then $\Delta\left(\theta^*\right)$ can be rewritten as

$$\Delta\left(\theta^*\right)=\frac{\mathbb{E}\left[\max\left(\theta,\theta^*\right)\right]\cdot\mathbb{E}\left[\min\left(\theta,\theta^*\right)\right]}{\theta^*}. \tag{C.69}$$

In turn, it suffices for us to prove that, when $\mathbb{E}\left(\theta\right)=1$, we always have $\mathbb{E}\left[\max\left(\theta,\theta^*\right)\right]\cdot\mathbb{E}\left[\min\left(\theta,\theta^*\right)\right]\leq\theta^*$.

Note that

$$\mathbb{E}\left[\max\left(\theta,\theta^*\right)\right]=\underbrace{\int_{\theta_{\min}}^{\theta^*}\theta^* dF}_{a_1}+\underbrace{\int_{\theta^*}^{\theta_{\max}}\theta dF}_{a_2}, \tag{C.70}$$

$$\mathbb{E}\left[\min\left(\theta,\theta^*\right)\right]=\underbrace{\int_{\theta_{\min}}^{\theta^*}\theta dF}_{b_1}+\underbrace{\int_{\theta^*}^{\theta_{\max}}\theta^* dF}_{b_2}. \tag{C.71}$$

and thus by definition, we know that

$$a_1 + b_2 = \theta^*, \tag{C.72}$$
$$a_2 + b_1 = \mathbb{E}(\theta). \tag{C.73}$$

Since we have normalized such that $\mathbb{E}(\theta) = 1$, it remains for us to prove that

$$(a_1 + a_2)(b_1 + b_2) \leq (a_1 + b_2)(a_2 + b_1), \tag{C.74}$$

which is equivalent to

$$(a_1 - b_1)(b_2 - a_2) \leq 0. \tag{C.75}$$

Note that

$$a_1 - b_1 = \int_{\theta_{\min}}^{\theta^*} (\theta^* - \theta)\, dF \geq 0, \tag{C.76}$$
$$a_2 - b_2 = \int_{\theta^*}^{\theta_{\max}} (\theta - \theta^*)\, dF \geq 0. \tag{C.77}$$

Therefore it is always held that

$$(a_1 - b_1)(b_2 - a_2) \leq 0, \tag{C.78}$$

where the equality holds iff $a_1 - b_1 = 0$, or $a_2 - b_2 = 0$, which respectively corresponds to $\theta^* = \theta_{\min}$ or $\theta^* = \theta_{\max}$.

Proof of Proposition 6: Since $W_{jt} = \alpha_j^l \frac{Y_{jt}}{L_{jt}}$, and $\frac{Y_{jt}}{W_{jt}} = \frac{\gamma_{jt} C_t}{W_t}$, and we have proved previously that

$$C_t = D(\theta_t^*) X_t = \frac{W_t}{\psi_t} \Delta_t. \tag{C.79}$$

where $\psi_t \equiv \frac{\psi}{(L_t/L)^{\nu}}$, and $\Delta_t \equiv \frac{D(\theta_t^*)\Gamma(\theta_t^*)}{(\varepsilon + \theta_t^*)}$. Then we know that

$$L_{jt} = \alpha_j^l \frac{Y_{jt}}{W_{jt}} = \frac{\alpha_j^l \gamma_{jt} C_t}{W_t} = \frac{\alpha_j^l \gamma_{jt}}{\psi_t} \Delta_t, \tag{C.80}$$

and thus

$$\gamma_{jt} = \psi_t \frac{\widetilde{L}_{jt}}{\Delta_t}, \tag{C.81}$$

where $\widetilde{L}_{jt} \equiv L_{jt}/\alpha_j^l$.

Substituting equation (C.81) into (C.63) yields

$$\sum_{i=1}^{N} \widetilde{L}_{jt} \xi_{ijt}^z = (1 - g_{jt}) \widetilde{L}_{jt} - \frac{\varphi_j \Delta_t}{\psi_t} - \beta \sum_{i=1}^{N} \mathbb{E}_t \alpha_i^s \omega_{ij} \widetilde{L}_{i,t+1}, \tag{C.82}$$

Meanwhile, equation (C.60) can be rewritten as

$$\widetilde{L}_{jt} \xi_{ijt}^z = \beta \mathbb{E}_t \left\{ \frac{\psi_{t+1}}{\psi_t} \left[\delta_i \alpha_i^k \phi_{ij} \widetilde{L}_{i,t+1} + (1 - \delta_i)\, \xi_{ij,t+1}^z \widetilde{L}_{j,t+1} \right] \right\}. \tag{C.83}$$

Furthermore, when $\delta_j = \delta$ for all $j \in \mathbf{N}$, combining equation (C.82) and (C.83) yields

$$
\begin{aligned}
(1-g_{jt})\widetilde{L}_{jt} &= \frac{\varphi_j}{\psi_t}\left(\Delta_t-\beta(1-\delta)\mathbb{E}_t\Delta_{t+1}\right)+\beta\mathbb{E}_t\frac{\psi_{t+1}}{\psi_t}(1-\delta)(1-g_{j,t+1})\widetilde{L}_{j,t+1} \\
&+\beta\mathbb{E}_t\sum_{i=1}^{N}\left(\alpha_i^s\omega_{ij}+\frac{\psi_{t+1}}{\psi_t}\delta\alpha_i^k\phi_{ij}\right)\widetilde{L}_{i,t+1} \\
&-\beta^2(1-\delta)\mathbb{E}_t\sum_{i=1}^{N}\mathbb{E}_t\frac{\psi_{t+1}}{\psi_t}\alpha_i^s\omega_{ij}\widetilde{L}_{i,t+2}. \qquad \text{(C.84)}
\end{aligned}
$$

Writing the above equation in a more compact way yields

$$
\left(\mathbf{E}-\mathbf{Diag}(\mathbf{g}_t)\right)\widetilde{\mathbf{L}}_t=\frac{\Delta_t-\beta(1-\delta)\mathbb{E}_t\Delta_{t+1}}{\psi_t}\boldsymbol{\varphi}+\boldsymbol{\rho}_{1t}\mathbb{E}_t\widetilde{\mathbf{L}}_{t+1}-\boldsymbol{\rho}_{2t}\mathbb{E}_t\widetilde{\mathbf{L}}_{t+2}, \qquad \text{(C.85)}
$$

where

$$
\begin{aligned}
\boldsymbol{\rho}_{1t} &\equiv \beta\left[(1-\delta)\frac{\psi_{t+1}}{\psi_t}\left(\mathbf{E}-\mathbf{Diag}(\mathbf{g}_{t+1})\right)+\delta\frac{\psi_{t+1}}{\psi_t}\mathbf{Diag}\left(\boldsymbol{\alpha}^k\right)\boldsymbol{\Phi}+\mathbf{Diag}(\boldsymbol{\alpha}^s)\boldsymbol{\Omega}\right]', \qquad \text{(C.86)} \\
\boldsymbol{\rho}_{2t} &\equiv \beta^2(1-\delta)\frac{\psi_{t+1}}{\psi_t}\left(\mathbf{Diag}(\boldsymbol{\alpha}^s)\boldsymbol{\Omega}\right)', \qquad \text{(C.87)}
\end{aligned}
$$

where $\mathbf{E}$ denotes $N\times N$ identity matrix, $\boldsymbol{\Omega}$ the input-output matrix with a typical element ω_{ij}.

Proof of Proposition 7: Denote $r\equiv 1/\beta-1$. Then equation (18) implies that the liquidity premium in steady state is given by

$$
\Gamma(\theta^*)=(1+\pi)(1+r), \qquad \text{(C.88)}
$$

Moreover, in steady state, equation (64) implies that

$$
\widetilde{\mathbf{L}}=\left[\frac{1-\beta(1-\delta)}{\psi}\Delta\right]\left[\mathbf{E}-\mathbf{Diag}(\mathbf{g})-\boldsymbol{\rho}_1+\boldsymbol{\rho}_2\right]^{-1}\boldsymbol{\varphi}. \qquad \text{(C.89)}
$$

Now we characterize $\left\{\gamma_j,\xi_{ij}^s,\xi_{ij}^z\right\}_{i,j\in\mathbf{N}}$. In steady state, $\gamma_{jt}=\gamma_j$, $\xi_{ijt}^z=\xi_{ij}^z$, and therefore we have

$$
\begin{aligned}
\frac{C_j}{Y_j} &= \xi_j^c\equiv\frac{\varphi_j}{\gamma_j}, \qquad \text{(C.90)} \\
\frac{S_{ij}}{Y_j} &= \xi_{ij}^s\equiv\beta\alpha_i^s\omega_{ij}\frac{\gamma_i}{\gamma_j}. \qquad \text{(C.91)}
\end{aligned}
$$

Besides, equation (C.60) immediately implies that

$$
\frac{Z_{ij}}{Y_j}=\xi_{ij}^z=\frac{\alpha_i^k\delta_i\phi_{ij}}{r+\delta_i}\frac{\gamma_i}{\gamma_j}, \qquad \text{(C.92)}
$$

It remains for us to solve $\boldsymbol{\gamma}$. Substituting equation (C.92) into (C.63) then yields

$$
(1-g_j)\gamma_j=\varphi_j+\beta\sum_{i=1}^{N}\left(\alpha_i^s\omega_{ij}+\frac{\delta_i\alpha_i^k\phi_{ij}}{1-\beta(1-\delta_i)}\right)\gamma_i, \qquad \text{(C.93)}
$$

and thus the vector of Dommar weight, $\boldsymbol{\gamma}$, is obtained as

$$\boldsymbol{\gamma} = \left(\mathbf{E} - \beta\widehat{\boldsymbol{\Omega}}'\right)^{-1} \boldsymbol{\varphi}. \tag{C.94}$$

Moreover, the generalized I-O matrix $\widehat{\boldsymbol{\Omega}}$ is given by

$$\widehat{\boldsymbol{\Omega}} \equiv \mathbf{Diag}\left(\boldsymbol{\alpha}^{\mathbf{s}}\right)\boldsymbol{\Omega} + \frac{1}{\beta}\mathbf{Diag}\left(\frac{\boldsymbol{\delta}}{\mathbf{r}+\boldsymbol{\delta}}\right)\mathbf{Diag}\left(\boldsymbol{\alpha}^{k}\right)\boldsymbol{\Phi} + \frac{1}{\beta}\mathbf{Diag}\left(\mathbf{g}\right), \tag{C.95}$$

and a typical element of $\mathbf{g}$ and $\frac{\boldsymbol{\delta}}{\mathbf{r}+\boldsymbol{\delta}}$ is respectively given by g_j and $\frac{\delta_j}{r+\delta_j}$. Additionally, $\mathbf{Diag}\left(\mathbf{g}\right)$ and $\mathbf{Diag}\left(\frac{\boldsymbol{\delta}}{\mathbf{r}+\boldsymbol{\delta}}\right)$ respectively denotes the diagonal matrix of $\mathbf{g}$ and $\frac{\boldsymbol{\delta}}{\mathbf{r}+\boldsymbol{\delta}}$. Additionally, some algebraic manipulation on equation (C.89) suggests that

$$\left[1 - \beta\left(1-\delta\right)\right]\left[\mathbf{E} - \mathbf{Diag}\left(\mathbf{g}\right) - \boldsymbol{\rho}_1 + \boldsymbol{\rho}_2\right]^{-1} = \boldsymbol{\gamma}, \tag{C.96}$$

where $\boldsymbol{\gamma}$ is given by equation (C.94).

Then the vector of the modified sectoral labor can be rewritten as

$$\widetilde{\mathbf{L}} = \frac{\gamma}{\psi}\Delta, \tag{C.97}$$

In turn, the sectoral labor is obtained as

$$L_j = \frac{\alpha_j^l \gamma_j}{\psi}\Delta. \tag{C.98}$$

Alternatively, equation (C.81) implies that $L_j = \frac{\alpha_j^l \gamma_j}{\psi}\Delta$ as in the above equation.

Now we characterize the sectoral output and sectoral capital. To begin with, the law of motion of sectoral capital in equation (32) implies that

$$I_i = \delta_i K_i, \tag{C.99}$$

and thus

$$\ln I_i = \ln K_i + \ln \delta_i. \tag{C.100}$$

Meanwhile, equation (33) implies

$$\ln I_i = \ln \epsilon_i + \sum_{j=1}^{N} \phi_{ij} \ln Z_{ij}. \tag{C.101}$$

Moreover, equation (C.92) can be rewritten as

$$\ln Z_{ij} = \ln \xi_{ij} + \ln Y_j. \tag{C.102}$$

where $\xi_{ij} = \frac{\alpha_i^k \delta_i \phi_{ij}}{r+\delta_i}\frac{\gamma_i}{\gamma_j}$. Then we know that

$$\ln I_i = \ln K_i + \ln \delta_i = \ln \epsilon_i + \sum_{j=1}^{N} \phi_{ij} \ln \xi_{ij} + \sum_{j=1}^{N} \phi_{ij} \ln Y_j, \tag{C.103}$$

and thus

$$\ln K_i = \ln c_i^k + \sum_{j=1}^{N} \phi_{ij} \ln Y_j, \tag{C.104}$$

where

$$\ln c_i^k \equiv \ln \epsilon_i + \sum_{j=1}^{N} \phi_{ij} \ln \xi_{ij} - \ln \delta_i, \tag{C.105}$$

or equivalently,

$$c_i^k \equiv \frac{\epsilon_i}{\delta_i} \prod_{j=1}^{N} \xi_{ij}^{\phi_{ij}}. \tag{C.106}$$

Furthermore, taking log on both sides of equation (31) yields

$$\ln Y_i = \ln A_i + \alpha_i^k \ln K_i + \alpha_i^l \ln L_i + \alpha_i^s \sum_{j=1}^{N} \omega_{ij} \ln S_{ij}, \tag{C.107}$$

Meanwhile, we have proved that

$$\ln S_{ij} = \ln \beta + \ln \left(\alpha_i^s \omega_{ij} \frac{\gamma_i}{\gamma_j} \right) + \ln Y_j, \tag{C.108}$$

and thus

$$\ln Y_i = \ln c_i + \alpha_i^k \ln K_i + \alpha_i^s \sum_{j=1}^{N} \omega_{ij} \ln Y_j, \tag{C.109}$$

where

$$\ln c_i \equiv \ln A_i + \alpha_i^l \ln L_i + \alpha_i^s \ln \beta + \alpha_i^s \sum_{j=1}^{N} \omega_{ij} \ln \left(\alpha_i^s \omega_{ij} \frac{\gamma_i}{\gamma_j} \right), \tag{C.110}$$

or equivalently,

$$c_i \equiv A_i L_i^{\alpha_i^l} \beta^{\alpha_i^s} \prod_{j=1}^{N} \left(\alpha_i^s \omega_{ij} \frac{\gamma_i}{\gamma_j} \right)^{\alpha_i^s \omega_{ij}}. \tag{C.111}$$

Combining equation (C.104) and (C.109) then gives

$$\ln Y_i = \ln c_i^y + \sum_{j=1}^{N} \left(\alpha_i^k \phi_{ij} + \alpha_i^s \omega_{ij} \right) \ln Y_j, \tag{C.112}$$

where

$$c_i^y \equiv \left(c_i^k \right)^{\alpha_i^k} c_i = A_i \left(\frac{\epsilon_i}{\delta_i} \prod_{j=1}^{N} \chi_{ij}^{\phi_{ij}} \right)^{\alpha_i^k} (L_i)^{\alpha_i^l} \prod_{j=1}^{N} \left(\beta \alpha_i^s \omega_{ij} \frac{\gamma_i}{\gamma_j} \right)^{\alpha_i^s \omega_{ij}}, \tag{C.113}$$

and L_i is given by equation (C.98). Then writing equation (C.112) in a more compact way yields

$$\ln \mathbf{Y} = \ln \mathbf{c}^y + \left[\mathbf{Diag} \left(\boldsymbol{\alpha}^k \right) \boldsymbol{\Phi} + \mathbf{Diag} \left(\boldsymbol{\alpha}^s \right) \boldsymbol{\Omega} \right] \ln \mathbf{Y}, \tag{C.114}$$

which immediately gives the sectoral output as

$$\ln \mathbf{Y} = \left[\mathbf{E} - \mathbf{Diag} \left(\boldsymbol{\alpha}^k \right) \boldsymbol{\Phi} - \mathbf{Diag} \left(\boldsymbol{\alpha}^s \right) \boldsymbol{\Omega} \right]^{-1} \ln \mathbf{c}^y, \tag{C.115}$$

with $\ln \mathbf{Y} \equiv [\ln Y_1, ..., \ln Y_N]'$ and $\mathbf{c}^y$ is given by

$$\begin{aligned}
\ln \mathbf{c}^y &= \mathbf{Diag} \left(\boldsymbol{\alpha}^k \right) \ln \mathbf{c}^k + \mathbf{Diag} \left(\boldsymbol{\alpha}^s \right) \ln \mathbf{c}^s + \ln \mathbf{A} + \mathbf{Diag} \left(\boldsymbol{\alpha}^l \right) \ln \mathbf{L}, \\
\ln \mathbf{c}^k &= \mathbf{Diag} \left(\ln \left(\boldsymbol{\Xi}^z \right) \boldsymbol{\Phi}' \right) + \ln \boldsymbol{\epsilon} - \ln \boldsymbol{\delta}, \\
\ln \mathbf{c}^s &= \mathbf{Diag} \left(\ln \left(\boldsymbol{\Xi}^s \right) \boldsymbol{\Phi}' \right),
\end{aligned}$$

where a typical element of $\boldsymbol{\Xi}^z$ and $\boldsymbol{\Xi}^s$ is respectively given by

$$\begin{aligned} \xi_{ij}^z &= \frac{Z_{ij}}{Y_j} = \frac{\alpha_i^k \delta_i \phi_{ij}}{r+\delta_i} \frac{\gamma_i}{\gamma_j}, && \text{(C.116)} \\ \xi_{ij}^s &= \frac{S_{ij}}{Y_j} = \frac{\alpha_i^s \omega_{ij}}{r+1} \frac{\gamma_i}{\gamma_j}. && \text{(C.117)} \end{aligned}$$

In turn, writing equation (C.104) in a compact way gives

$$\ln \mathbf{K} = \ln \mathbf{c}^k + \boldsymbol{\Phi} \ln \mathbf{Y}, \quad \text{(C.118)}$$

where $\ln \mathbf{K} \equiv [\ln K_1, ..., \ln K_N]'$. Then we immediately sectoral investment as $I_j = \delta_j K_j$, and thus

$$\ln \mathbf{I} = \ln \boldsymbol{\delta} + \ln \mathbf{K}. \quad \text{(C.119)}$$

Similarly,

$$\ln \mathbf{C} = \ln \mathbf{Y} + \ln \boldsymbol{\varphi} - \ln \boldsymbol{\gamma}.$$

Finally, we analyze how factor prices are determined. First, given that we have obtained $\{Y_j, S_{ij}, K_j, L_j\}$, Proposition 4 implies that, in steady state, the factor prices are given by

$$\begin{aligned} R_j^k &= \alpha_j^k \frac{Y_j}{K_j}, && \text{(C.120)} \\ R_{ij}^s &= \alpha_i^s \omega_{ij} \frac{Y_i}{S_{ij}}, && \text{(C.121)} \\ W_j &= \alpha_j^l \frac{Y_j}{L_j}, && \text{(C.122)} \end{aligned}$$

Moreover, the normalization of the price index of the final goods is given by

$$Q \equiv \prod_{j=1}^{N} \left(\frac{Q_j}{\varphi_j} \right)^{\varphi_j} = 1, \quad \text{(C.123)}$$

where $Q_j = W/W_j$. Therefore the wage rate is obtained as

$$W = \prod_{j=1}^{N} \left(\varphi_j W_j \right)^{\varphi_j}, \quad \text{(C.124)}$$

and thus

$$\ln W = \boldsymbol{\varphi}' \left(\ln \boldsymbol{\varphi} + \ln \mathbf{W} \right).$$

Since we have obtained (W, W_j), we easily obtain the price of goods produced in sector j as

$$Q_j = \frac{W}{W_j}. \quad \text{(C.125)}$$

Writing in a more way for R_j^k, W_j, Q_j yields

$$\begin{aligned} \ln \mathbf{R}^k &= \ln \boldsymbol{\alpha}^k + \ln \mathbf{Y} - \ln \mathbf{K}, && \text{(C.126)} \\ \ln \mathbf{W} &= \ln \boldsymbol{\alpha}^l + \ln \mathbf{Y} - \ln \mathbf{L}, && \text{(C.127)} \\ \ln \mathbf{Q} &= \boldsymbol{\varphi}' \left(\ln \boldsymbol{\varphi} + \ln \mathbf{W} \right) \mathbf{1} - \ln \mathbf{W}. && \text{(C.128)} \end{aligned}$$

Proof of Corollary x.x: In complete market, $\tau^w = 0$, and $\Delta = 1$. Since $\widetilde{\mathbf{L}} = \frac{\gamma}{\psi}\Delta$, immediately we have

$$\ln \widetilde{\mathbf{L}} - \ln \widetilde{\mathbf{L}}^* = \ln \mathbf{L} - \ln \mathbf{L}^* - \tau^w, \tag{C.129}$$

where $\tau^w = -\ln \Delta(\theta^*)$ denotes the labor wedge, where θ^* is determined by equation (67), and where the superscript $*$ denotes variables under complete market (except the cutoff θ^*). In turn, equation (C.113) implies that

$$\ln c_i^y - \ln c_i^{y*} = \alpha_i^l (\ln L_i - \ln L_i^*). \tag{C.130}$$

In turn, Proposition 7 suggests that **[To be revised]**

$$\left[\mathbf{E} - \mathbf{Diag}\left(\boldsymbol{\alpha}^k\right)\boldsymbol{\Phi} - \mathbf{Diag}\left(\boldsymbol{\alpha}^s\right)\boldsymbol{\Omega}\right]^{-1}\alpha,$$

and then we immediately know that

$$\begin{aligned}
\ln \mathbf{K} - \ln \mathbf{K}^* &= \boldsymbol{\Phi}(\ln \mathbf{Y} - \ln \mathbf{Y}^*), && \text{(C.131)}\\
\ln \mathbf{I} - \ln \mathbf{I}^* &= \boldsymbol{\Phi}(\ln \mathbf{Y} - \ln \mathbf{Y}^*), && \text{(C.132)}\\
\ln \mathbf{C} - \ln \mathbf{C}^* &= \ln \mathbf{Y} - \ln \mathbf{Y}^*, && \text{(C.133)}\\
\ln C - \ln C^* &= \boldsymbol{\varphi}'(\ln \mathbf{Y} - \ln \mathbf{Y}^*). && \text{(C.134)}
\end{aligned}$$

where the final consumption is given by $C = \prod_{j=1}^{N} C_j^{\varphi_j}$.

D Dynamical System

Dynamical System

$$C_{jt} = \frac{\varphi_j}{Q_{jt}} D_t X_t, \tag{D.1}$$

$$\frac{M_{t+1}}{P_t} = H(\theta_t^*) X_t, \tag{D.2}$$

$$(\varepsilon + \theta_t^*) X_t = \frac{\Gamma_t}{\Lambda_t}, \tag{D.3}$$

$$1 = \beta \mathbb{E}_t \frac{\Lambda_{t+1}}{\Lambda_t} \frac{P_t}{P_{t+1}} \Gamma_t, \tag{D.4}$$

$$Q_{jt} = \beta \mathbb{E}_t \frac{\Lambda_{t+1}}{\Lambda_t} Q_{i,t+1} R^s_{ij,t+1}, \tag{D.5}$$

$$\Lambda_t = \frac{\psi_t}{W_t}$$

$$Y_{it} = A_{it} K_{it}^{\alpha_i^k} L_{it}^{\alpha_i^l} \left(\prod_{j=1}^{N} S_{ijt}^{\omega_{ij}}\right)^{\alpha_i^s}, \tag{D.6}$$

$$K_{i,t+1} = K_{it}^{1-\delta_i} (I_{it}/\delta_i)^{\delta_i}, \tag{D.7}$$

$$I_{it} = \epsilon_{it} \prod_{j=1}^{N} Z_{ijt}^{\phi_{ij}}, \tag{D.8}$$

$$R_{jt}^{k} = \alpha_{j}^{k} \frac{Y_{jt}}{K_{jt}}, \tag{D.9}$$

$$W_{jt} = \alpha_{j}^{l} \frac{Y_{jt}}{L_{jt}}, \tag{D.10}$$

$$R_{ijt}^{s} = \alpha_{i}^{s} \omega_{ij} \frac{Y_{it}}{S_{ijt}}, \tag{D.11}$$

$$W_{t} = Q_{jt} W_{jt}, \tag{D.12}$$

$$C_{jt} + \sum_{i=1}^{N} S_{ij,t+1} + \sum_{i=1}^{N} Z_{ijt} + T_{jt} = Y_{jt}, \tag{D.13}$$

$$\psi_{t} = \frac{\psi}{(L_{t}/L)^{\nu}}, \tag{D.14}$$

$$\Gamma_{t} = \mathbb{E} \max \left(\varepsilon + \theta, \varepsilon + \theta_{t}^{*} \right), \tag{D.15}$$

$$D_{t} = \mathbb{E} \min \left(\varepsilon + \theta, \varepsilon + \theta_{t}^{*} \right) = 1 + \varepsilon + \theta_{t}^{*} - \Gamma_{t}, \tag{D.16}$$

$$H_{t} = 1 - D_{t}, \tag{D.17}$$

$$\bar{M}_{t+1} = M_{t+1} = M_{t} + \tau_{t}. \tag{D.18}$$

Log-linearization

$$\widehat{C}_{jt} = \widehat{D}_{t} + \widehat{X}_{t} - \widehat{Q}_{jt}, \tag{D.19}$$

$$\widehat{M}_{t+1} - \widehat{P}_{t} = \widehat{H}_{t} + \widehat{X}_{t}, \tag{D.20}$$

$$\frac{\theta^{*}}{\varepsilon + \theta^{*}} \widehat{\theta}_{t}^{*} + \widehat{X}_{t} = \widehat{\Gamma}_{t} - \widehat{\Lambda}_{t}, \tag{D.21}$$

$$\widehat{P}_{t+1} - \widehat{P}_{t} = \widehat{\Gamma}_{t} + \left(\widehat{\Lambda}_{t+1} - \widehat{\Lambda}_{t} \right), \tag{D.22}$$

$$\widehat{S}_{ij,t+1} = -\widehat{Q}_{jt} + \left(\widehat{Q}_{i,t+1} + \widehat{Y}_{i,t+1} \right) + \left(\widehat{\Lambda}_{t+1} - \widehat{\Lambda}_{t} \right), \tag{D.23}$$

$$\begin{aligned} \widehat{Z}_{ijt} \quad = \quad & -\widehat{Q}_{jt} + \left(\widehat{\Lambda}_{t+1} - \widehat{\Lambda}_{t} \right) \\ & + \beta \left(r + \delta_{i} \right) \left(\widehat{Q}_{i,t+1} + \widehat{Y}_{i,t+1} \right) + \beta \left(1 - \delta_{i} \right) \left(\widehat{Q}_{j,t+1} + \widehat{Z}_{ij,t+1} \right), \end{aligned} \tag{D.24}$$

$$\widehat{\Lambda}_{t} = \widehat{\psi}_{t} - \widehat{W}_{t},$$

$$\widehat{Y}_{it} = \widehat{A}_{it} + \alpha_{i}^{k} \widehat{K}_{it} + \alpha_{i}^{l} \widehat{L}_{it} + \alpha_{i}^{s} \sum_{j=1}^{N} \omega_{ij} \widehat{S}_{ijt}, \tag{D.25}$$

$$\widehat{K}_{i,t+1} = \left(1 - \delta_{i} \right) \widehat{K}_{it} + \delta_{i} \widehat{I}_{it}, \tag{D.26}$$

$$\widehat{I}_{it} = \widehat{\epsilon}_{it} + \sum_{j=1}^{N} \phi_{ij} \widehat{Z}_{ijt}, \tag{D.27}$$

$$\widehat{W}_{jt} = \widehat{Y}_{jt} - \widehat{L}_{jt}, \tag{D.28}$$

$$\widehat{R}_{jt}^{k} = \widehat{Y}_{jt} - \widehat{K}_{jt}, \tag{D.29}$$

$$\widehat{R}^s_{ijt} = \widehat{Y}_{jt} - \widehat{S}_{ijt}, \tag{D.30}$$

$$\widehat{W}_t = \widehat{Q}_{jt} + \widehat{W}_{jt}, \tag{D.31}$$

$$\xi^c_j \widehat{C}_{jt} + \sum_{i=1}^{N} \xi^s_{ij} \widehat{S}_{ij,t+1} + \sum_{i=1}^{N} \xi^z_{ij} \widehat{Z}_{ijt} + g_j \widehat{T}_{jt} = \widehat{Y}_{jt}, \tag{D.32}$$

$$\widehat{\psi}_t = -\nu \widehat{L}_t, \tag{D.33}$$

$$\widehat{M}_{t+1} = \rho_m \widehat{M}_t + \varepsilon^m_t, \tag{D.34}$$

$$\widehat{\Gamma}_t = \frac{F\left(\theta^*\right)\theta^*}{\Gamma} \widehat{\theta}^*_t, \tag{D.35}$$

$$\widehat{D}_t = \frac{\theta^*}{D} \widehat{\theta}^*_t - \frac{\Gamma}{D} \widehat{\Gamma}_t, \tag{D.36}$$

$$\widehat{H}_t = -\frac{D}{H} \widehat{D}_t, \tag{D.37}$$

where $\Gamma = (1+\pi)(1+r)$, $r = 1/\beta - 1$, $D = 1 + \varepsilon + \theta^* - \Gamma$, $H = 1 - D$, and

$$\begin{aligned}
\xi^c_j &= \frac{C_j}{Y_j} = \frac{\varphi_j}{\gamma_j}, \\
\xi^z_{ij} &= \frac{Z_{ij}}{Y_j} = \alpha^k_i \phi_{ij} \frac{\delta_i}{r+\delta_i} \frac{\gamma_i}{\gamma_j}, \\
\xi^s_{ij} &= \frac{S_{ij}}{Y_j} = \alpha^s_i \omega_{ij} \frac{1}{r+1} \frac{\gamma_i}{\gamma_j}, \\
g_j &= \frac{T_j}{Y_j}.
\end{aligned}$$