Staff Working Paper/Document de travail du personnel — 2021-46
Last updated: September 29, 2021

More Money for Some: The Redistributive Effects of Open Market Operations

by Christian Bustamante

Financial Stability Department
Bank of Canada, Ottawa, Ontario, Canada K1A 0G9

CBustamante@bank-banque-canada.ca

ISSN 1701-9397

Acknowledgements

This paper was previously circulated as "More Money for Some: Monetary Policy Meets a Rich and Persistent Household Wealth Distribution." I thank Aubhik Khan and Julia K. Thomas for their invaluable comments. I also thank Jonathan Chiu, Ruben Hipp, Heejeong Kim, Miguel Molico, Thomas Pugh, and session participants at the 2018 North American Summer Meeting of the Econometric Society at UC Davis, 2017 Midwest Macro Meetings at University of Pittsburgh, as well as workshop participants at Ohio State, for comments and suggestions that have improved this paper.

Abstract

Using a general equilibrium search-theoretic model of money, I study the distributional effects of open market operations. In my model, heterogeneous agents trade bilaterally among themselves in a frictional market and save using cash and illiquid short-term nominal government bonds. Wealth effects generate slow adjustments in agents' portfolios following their trading activity in decentralized markets, giving rise to a persistent and nondegenerate distribution of assets. The model reproduces the distribution of asset levels and portfolios across households observed in the data, which is crucial to quantitatively assess the incidence of monetary policy changes at the individual level. I find that an open market operation targeting a higher nominal interest rate requires increasing the relative supply of bonds, raising the ability of agents to self-insure against idiosyncratic shocks. As a result, in the long run, inequality falls, and the inefficiencies in decentralized trading shrink. This leads agents that are relatively poor and more liquidity-constrained to benefit the most by increasing their consumption and welfare.

Topics: Inflation and prices, Monetary policy, Monetary policy implementation, Monetary policy transmission

JEL codes: E21, E32, E52

1 Introduction

Open market operations (OMOs) are among the most common instruments that central banks use to implement monetary policy. With them, central banks exchange cash for short-term government bonds. However, there is little understanding about how these interventions can disproportionally affect households with different levels of wealth and portfolio compositions. This paper studies the transmission mechanism of monetary policy via OMOs when accounting for the observed persistent differences in asset levels and portfolios across households. Further, it examines the distributional consequences of such an implementation and their interaction with the transmission mechanism of monetary policy itself.

Rather than assuming that the economy adjusts instantly and without any cost to a certain targeted interest rate, I explicitly model the transmission mechanism from OMOs to liquidity and interest rates and then to the rest of the economic activity. Given that not all agents are receiving the same extra liquidity after an OMO, I conduct the analysis in a context in which changes in the state of the economy (persistent distribution of assets) are fundamental in determining the transmission of monetary policy because of the distributional consequences of OMOs. In this spirit, this paper touches on several questions related to the interplay of OMOs and heterogeneity in asset holdings: are these operations neutral in the long run so that changes in monetary variables have no impact on the real economic activity? Does monetary policy affect households in an environment where individual gains vary due to the existing endogenous heterogeneity? How much does the effectiveness of an OMO depend on the level of trend inflation? What does that imply for the optimal conduct of monetary policy?

I build a search-theoretic model of money where agents trade sequentially in a frictionless centralized market (CM) and in a frictional decentralized market (DM) that generates the necessity of using money for transactional purposes. In the DM, agents are randomly matched in bilateral meetings. Agents do not have access to credit but, besides holding money, can save in illiquid nominal short-term government bonds. These bonds are illiquid, in the sense that they cannot be used in decentralized trading. However, there is a positive demand for bonds because agents can use them to partially self-insure against idiosyncratic shocks and, therefore, smooth consumption in the CM. A unified fiscal and monetary authority controls the supply of money and nominal bonds. By managing the size and composition of public liquidity, the monetary authority affects the real economy. Such changes in the relative supply of money and government bonds are implemented by means of OMOs.

The extent to which agents are disproportionally affected by an OMO depends on the size and composition of their portfolios. For example, after an OMO in which the central bank buys bonds by injecting more money, agents with large levels of bond holdings are the ones receiving most of that additional liquidity. This type of intervention may have significantly different implications than the monetary injections via lump-sum transfers that have been traditionally studied in the literature and that give every agent the same extra liquidity independent of their location in the distribution of wealth.

Distributional effects are crucial to study the implications of OMOs. In the absence of changes in agents' budget sets, OMOs become irrelevant and do not affect the real economy. As discussed by Wallace (1981) and Lucas (1984), if government taxes or transfers offset the distributional changes associated with OMOs, budget sets remain identical across the distribution of agents, and therefore, monetary policy becomes neutral. An environment with frictionless asset markets and with a representative agent would exhibit this particular type of Ricardian equivalence. By endogenously modeling the heterogeneity in asset holdings using an incomplete market setup, my model breaks this irrelevance result and allows us to investigate the consequences of OMOs in a richer and more realistic setting.

The presence of an interest-bearing asset implies that the insurance role of money is attenuated in my model. Rather than holding large stocks of money, agents are more prone to reduce their usage of money, leaving it mainly for transactional purposes while relying on bonds as their main saving instrument. Of course, the extent to which agents show this behavior depends on the trade-off between the insurance value of money for DM transactions and the insurance role of bonds in the CM. The strength of this trade-off determines portfolios and depends on wealth. Hence, capturing the differences in portfolios along the distribution of wealth observed in the data is critical for determining how OMOs affect different agents. For this reason, I calibrate the model to reproduce the size and composition of households' portfolios in the Survey of Consumer Finances (SCF).

I find that a higher concentration of wealth leads to more inefficient decentralized trading outcomes and reduces total output. Given that money is the only asset used as a means of payment in bilateral meetings, poorer agents are characterized by having portfolios with more money and less illiquid assets. This model result is consistent with the data. However, despite money being an essential asset, it is the total wealth of both agents that determines the opportunity cost of sellers and the buyers' willingness to pay. I show that the more uneven the levels of wealth between buyers and sellers, the higher the wedge between buyers' marginal utility and sellers' marginal cost. Therefore, the more skewed the distribution of wealth is, the more prevalent the inefficiencies in bilateral meetings.

The long-run effects of OMOs also hinge on how they interact with the distribution of wealth. In contrast to lump-sum transfers, since OMOs swap bonds for money, agents with more bond holdings experience most of the direct effects associated with this policy. The effects of OMOs then depend on how those changes in liquidity spread out from the top of the wealth distribution to the rest of the economy. In this context, for my baseline calibration, I first show that targeting an increase in the nominal interest rate of 1 percentage point (p.p.) requires conducting a (permanent) OMO that increases the relative supply of government bonds held by the public by 20 percent. When this happens, it becomes easier for agents to self-insure against possible expenditure shocks when trading in the DM. The increased insurance opportunities reduce wealth inequality, and hence, boost total output by 0.21 percent following an attenuation of the inefficiencies present in decentralized trading. Furthermore, these effects are not linear. An OMO that aims to lower the nominal interest rate in the same magnitude reduces total output by 0.31 percent.

Second, I quantify the effect of OMOs on agents with different levels of wealth. When an OMO increases the relative supply of bonds, the poorest agents gain the most as they increase their total consumption, work less, and become wealthier. This increased level of wealth in the left end of the distribution allows poorer agents to obtain better terms of trade in bilateral meetings, increasing not only their individual welfare but also total output. For an increase in the nominal rate of 1 p.p., all agents below the median level of wealth experience higher consumption levels in both markets. The bulk of this effect is concentrated in agents in the bottom quartile of the wealth distribution, with increases in total consumption of almost 2.5 percent.

I also assess the role of trend inflation and the presence of asymmetries in the long run after an OMO. I find that, with lower trend inflation, an OMO that increases the provision of liquidity benefits principally the poorest agents. When the long-run value of inflation is lower, nominal assets lose their value less rapidly, and agents spend their assets at a slower pace. This, in turn, implies that agents trade at lower bilateral prices so that the idiosyncratic risk associated with DM meetings falls and agents have less precautionary saving motives. In this context, the size of an OMO needs to be larger to reach the same targeted rate, which benefits the poorest and more liquidity-constrained agents.

In my model, a zero-inflation and zero-nominal-interest-rate policy maximizes output and welfare for a given level of provision of public liquidity. Akin to the Friedman rule, this policy ends up being optimal because nominal assets lose their value more slowly for lower levels of inflation. As a result, wealth is distributed more evenly, and agents trade at lower and less disperse prices. However, despite this being the case, the output-maximizing policy does not eliminate the inefficiencies associated with decentralized trading. This happens as agents trade at a suboptimal level as long as there are differences in wealth between buyers and sellers. The continuous presence of wealth effects and uninsurable risk guarantees that this is the case even for the output-maximizing policy. By the same token, the welfare cost of inflation rises alongside the dispersion in agents' wealth. I estimate that the welfare cost of a 10 percent inflation rate, relative to the baseline of no inflation, is 1.67 percent in terms of consumption.

Related Literature

This paper is related to several strands of the literature. First, the model economy builds on a vast literature of search-theoretic models of monetary exchange that includes papers such as Kiyotaki and Wright (1989), Trejos and Wright (1995), Shi (1997), Lagos and Wright (2005) and Rocheteau and Wright (2005).[1] These papers do not study OMOs but the equilibrium properties of economies where money is essential for trading. To do so, they either have to restrict the divisibility of assets or goods exchanged in these markets, as in Kiyotaki and Wright (1989) and Trejos and Wright (1995), or assume a certain mechanism that eliminates the heterogeneity intrinsically generated by decentralized trading. Shi (1997) assumes the existence of large families so that agents that are in principle different after trading bilaterally have access to a perfect risk-sharing instrument that collapses the distribution of money. Meanwhile, Lagos and Wright (2005) suppose that agents can participate in a CM immediately after being matched and exchanged in a bilateral meeting. The preferences of these agents in this Walrasian market are assumed to be quasi-linear. Consequently, any wealth effects that would operate under more general preferences are absent, and every agent decides to bring the same amount of money

[1] Lagos, Rocheteau and Wright (2017) present a comprehensive review of the literature about New Monetarist models, where they summarize the existing work on liquidity, monetary policy, and monetary and credit arrangements.

to the next DM. The distribution of money in the Lagos-Wright model, as well as in Shi (1997), is then degenerate. Their critical reason for trying to work with a degenerate distribution of money is to solve the agents' problem in a manageable manner. If the distribution is nondegenerate, agents must take into account the whole distribution and its evolution when making consumption and saving decisions. My paper departs from these assumptions and tackles the consequences of accounting for the persistent heterogeneity observed in the data.

Second, my model is also related to other papers that study the role of heterogeneity in search-theoretic models of money. Molico (2006); Chiu and Molico (2010); Chiu and Molico (2011); Menzio, Shi and Sun (2013); Rocheteau, Weill and Wong (2015); and Chiu and Molico (2020) address this topic. Some of these papers explore the redistributive consequences of different money injection mechanisms. However, they do so in environments with only one asset and where the distribution of assets does not determine the incidence of monetary policy at the individual level. The numerical method I developed to solve the stationary equilibrium may be of independent interest. As opposed to Chiu and Molico (2011), the method does not require simulating a large number of agents to then fit a distribution using kernel density estimation methods. Likewise, in contrast to Rocheteau, Weill and Wong (2015), my method is able to handle heterogeneity in both buyers and sellers so that we do not need to impose their types exogenously. My solution strategy relies on nonlinear methods and can accommodate more than one dimension in the agents' individual state space. I discuss the algorithm in detail in Appendix A.1.

Third, this paper contributes to the literature on OMOs and liquidity for the conduct of monetary policy.[2] Rocheteau, Wright and Xiao (2018) also study the effect of OMOs and characterize the different types of equilibria that can emerge under several types of market structures and levels of assets' liquidity. They show that when assets have different levels of pledgeability and hence liquidity, OMOs can have real effects. However, to ensure their model's analytical tractability, Rocheteau, Wright and Xiao (2018) assume quasi-linear preferences, and therefore eliminate any eventual persistent heterogeneity among agents, rendering their model money-neutral.

Alvarez, Atkeson and Edmond (2009) and Khan and Thomas (2015) study the role of monetary policy in economies with segmented markets where not all agents have access to asset markets. Kocherlakota (2003) examines the role of bonds and shows that having illiquid nominal risk-free bonds in monetary economies is essential, as they allow agents

[2]Some of the earlier work in this area includes papers such as Lucas (1972), Lucas (1984), Wallace (1981), Sargent and Smith (1987), Sargent and Wallace (1981), and Lucas (1990).

to engage in additional intertemporal exchanges of money. Similarly, Shi (2008) finds that there are efficiency gains of having this type of asset, as it provides an instrument for agents to (partially) self-insure against idiosyncratic liquidity shocks. He also argues that it is optimal for a government to impose legal restrictions on the liquidity of public debt. Geromichalos, Licari and Suárez-Lledó (2007) investigate the interaction between monetary policy and asset prices in a model with both money and real assets.

Williamson (2012) and Araujo and Ferraris (2020) discuss the effect of OMOs under different types of equilibria in an economy without persistent heterogeneity. They show that under a liquidity trap equilibrium, OMOs are not effective in the margin since the nominal interest rate is stuck at zero and total liquidity remains scarce. Sterk and Tenreyro (2018) explore the effect of OMOs and the transmission of monetary policy through durable consumption. As in my model, in their environment, wealth effects along the distribution crucially affect agents' asset accumulation decisions. Rocheteau and Rodriguez-Lopez (2014) assess the effects of an OMO that increases the public provision of real government bonds in an economy with over-the-counter markets, collateralized loans, and unemployment, and find that it reduces the private supply of liquidity and increases interest rates and unemployment. Carli and Gomis-Porqueras (2021) examine the role of unsecured credit and limited commitment for the conduct of monetary policy through OMOs.

The remainder of the paper is organized as follows. Section 2 outlines the model economy. Section 3 describes the calibration strategy. Later, Section 4 presents the properties of the stationary equilibrium. Section 5 discusses the long-run properties of implementing a change in the composition of publicly provided liquidity using OMOs, and Section 6 explores the mechanisms determining the welfare cost of inflation. Finally, Section 7 concludes.

2 Model

The model presented here is based in Lagos and Wright (2005) and Rocheteau, Wright and Xiao (2018), but with a few critical differences that will be discussed later. There is a contin-uum of agents of unitary mass. Time is discrete, agents live for infinite periods, and they

discount time at the rate $\beta \in (0, 1)$. Each period is divided into two markets that oper-ate sequentially: first, agents trade bilaterally in a decentralized market (DM), and later, they meet in a centralized market (CM). There is a unified fiscal and monetary authority that we will call government, which controls the supply of the two assets available to agents in this economy: fiat money and one-period nominal government bonds. Both of them are storable and perfectly divisible. Money has no intrinsic value, while government bonds represent claims of money in the next CM. Individual nominal money and bond holding will be normalized with respect to the beginning of the period money supply, M. Thus, if $\hat{m}$ and $\hat{a}$ are individual nominal money and bonds, respectively, then $m = \hat{m}/M$ and $a = \hat{a}/M$ denote individual relative money and bond holdings.

In the DM, agents may be randomly matched with others in bilateral meetings. Each agent specializes in producing a nonstorable differentiated good, c_d, that is "wanted" by other agents but not by that agent. This differentiated good can only be produced in the DM, and its technology of production is one-to-one on labor. The agents' period utility function is:

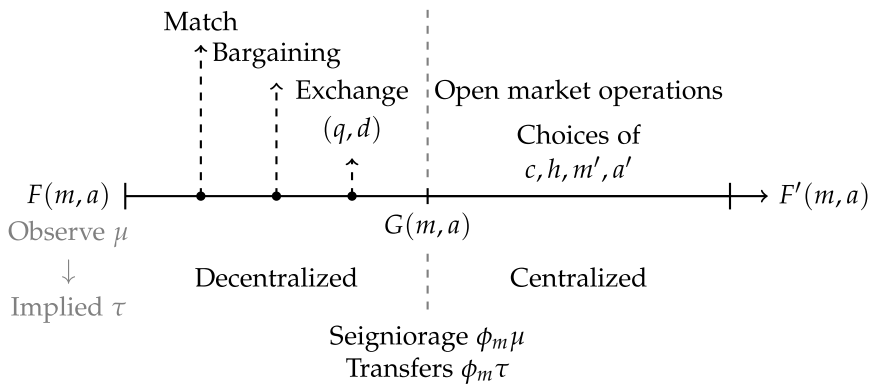
$$\mathcal{U}(c_d, h_d, c, h) = u(c_d) - v(h_d) + U(c, 1-h), \tag{1}$$

where u and v are, respectively, the utility function and cost function in the DM and U is the utility function in the CM. Here, c_d and h_d denote consumption and labor in the DM, while c and h are their counterparts in the CM (conversely, $1-h$ is leisure). The functions u and v satisfy: $u(0) = v(0) = 0$, $u' > 0$, $v' < 0$, $u'' < 0$, and $v'' \geq 0$; while U is twice continuously differentiable and strictly concave in both consumption and leisure. Unlike in Lagos and Wright (2005), here U is not restricted to be linear in h and, therefore, we will not necessarily have a degenerate distribution of assets.

When two agents are matched, we have two different types of meetings. If agent A wants what B produces but B does not want what A produces, we say we have a single-coincidence meeting where A is the buyer (consumer) and B is the seller (producer). We also have the possibility of not having any coincidence at all, where none of the matched agents like what the other produces.

Let α denote the probability of being matched, and let the probability of being matched as a buyer be $1/2$. When matched, the agents must determine the terms of trade between them: how much will be produced and at what price. Here, I assume that agents are anonymous, that their histories are private information, and that there is no record keeping. The only individual characteristics that are observable are asset holdings. However, bonds cannot be used in decentralized trading. These frictions prevent the existence of credit and make money essential as it is the only instrument they can use as a medium of exchange in meetings where barter is not an option (i.e., single-coincidence meetings). In this context, terms of trade are determined by a "take it or leave it" offer of the buyer to the seller.

Figure 1: Timing of Events



In the second subperiod, agents trade in a centralized (Walrasian) market. Agents have access to a production technology that transforms one unit of labor into one unit of a perishable homogeneous good. In this subperiod, agents can also adjust their asset holdings and receive lump-sum transfers from the government. The centralized market is also the time window used by the government to implement any change in its monetary policy involving OMOs. For example, if the government decides to increase its supply of money relative to government bonds, it will conduct a purchase of bonds that will, in turn, increase agents' money holdings. Of course, and this is a crucial point in this paper, agents entering the CM with different portfolios will be affected differently by this policy. Below, I present the model structure in more detail. See Figure 1 for an illustration of the exact timing of the model for a given period.

2.1 Government

The unified fiscal and monetary authority controls the supply of money, M, and nominal bonds, A. Let $K \equiv A/M$ denote the ratio of aggregate government bonds to aggregate money. Money grows at the constant rate μ^M, so $M' = (1+\mu^M)M$, while bonds do so at the rate μ^A. I denote the state vector characterizing monetary policy as $\theta \equiv (K, \mu)$ where $\mu \equiv (\mu^M, \mu^A)$. Stationarity requires A to grow at the same rate of M, so that $\mu^M = \mu^A = \overline{\mu}$ and K is a constant. I also assume that the government maintains a balanced budget. Hence, it finances its debt service and lump-sum transfers to agents with seigniorage and by issuing new one-period bonds. In this context, the government's balance is:

$$\phi_m K + T = \mu^M \phi_m + \left(1+\mu^A\right) \phi_m \phi_a K, \tag{2}$$

where ϕ_m and ϕ_a are the price of relative money and relative bonds, respectively. Note that the budget constraint in (2) is already expressed in terms of relative units.

Finally, government transfers (or taxes, if negative), are expressed in units of money. From (2), this implies that $T = \phi_m \tau$, where

$$\tau = \mu^M - \left[1 - \left(1 + \mu^A\right) \phi_a\right] K. \tag{3}$$

Note that the real interest rate on public debt is given by:

$$1 + r = \frac{1}{\phi_a}. \tag{4}$$

Also, by the Fischer equation, the nominal counterpart of r is $(1 + i) = (1 + r)(1 + \pi)$, where π denotes the inflation rate and is defined as $(1 + \pi) = (1 + \mu_{-1}^M)\phi_{m,-1}/\phi_m$. These definitions are useful for calibrating the model and interpreting its results.

2.2 Decentralized Market

The probability measures $F(m, a)$ and $G(m, a)$ summarize the distribution of agents over m and a for the DM and CM, respectively. These distributions are defined on the Borel algebra, $\mathcal{Z}$, generated by the open subsets of the space, $Z = \mathcal{M} \times \mathcal{A}$. Since the probability of being matched with someone with a particular level of money and bonds depends on how many of those agents are in the entire population, the probability distribution is the aggregate state variable. These distributions evolve according to $G = \Gamma_G(F, \theta)$ and $F' = \Gamma_F(G, \theta)$. Given this, and the policy implemented by the monetary authority, in equilibrium we have:

$$\int_{\mathcal{M} \times \mathcal{A}} m dF([dm \times da]) = 1 \tag{5}$$

$$\int_{\mathcal{M} \times \mathcal{A}} a dF([dm \times da]) = K \tag{6}$$

Let $\tilde{m}$ represent money holdings at the end of the DM and define $x \equiv \tilde{m} + a + \tau$ as wealth. Given that bonds are nominal, and that government transfers are also expressed in terms of money, x is the only relevant individual state variable for agents. Now, let $V(m, a; F, \theta)$ and $W(x; G, \theta)$ be the value functions at the beginning of the DM and CM, respectively. Since we are only considering single-coincidence meetings, we have no situations where barter is an option to conduct the exchange process. I also assume that, since bonds are illiquid, money is the only acceptable asset when exchanging in any of the goods markets. When two agents are randomly matched, one as a buyer and the other as a seller, they decide the terms of trade. I assume that the buyer makes a "take it or leave it" offer to the seller: the buyer offers to buy q of the differentiated good at the price d.

In a meeting where the buyer's portfolio is $z_b = (m_b, a_b)$ and the seller's is $z_s = (m_s, a_s)$, terms of trade are determined according to the following problem:

$$\max_{q,d} \; u(q) + W(m_b + a_b + \tau - d; G, \theta), \tag{7}$$

subject to the seller's participation constraint

$$-v(q) + W(m_s + a_s + \tau + d; G, \theta) \geq W(m_s + a_s + \tau; G, \theta), \tag{8}$$

to the law of motion of the distribution $G(m) = \Gamma_G(F(m), \theta)$, and to $0 \leq d \leq m_b$, $q \geq 0$. Note that the continuation values of buyers and sellers take into account their money and bond holdings when exiting the current DM and the transfer they will receive at the beginning of the next CM. For each meeting of type (m_b, a_b, m_s, a_s), and given an aggregate state $\{F(m,a), \theta\}$, we have that the terms of trade $q(z_b, z_s) \equiv q(m_b, a_b, m_s, a_s; F, \theta)$ and $d(z_b, z_s) \equiv d(m_b, a_b, m_s, a_s; F, \theta)$ solve the problem stated above.

In this context, the expected lifetime utility at the beginning of the DM of an agent with portfolio (m, a)—i.e., before knowing if they are matched or not, and before knowing what their role in an eventual match would be—is given by the following functional equation:

$$\begin{aligned} V(m, a; F, \theta) = &\frac{\alpha}{2} \int_{\mathcal{M} \times \mathcal{A}} \{u(q(z, z_s)) + W(m - d(z, z_s) + a + \tau; G, \theta)\} F(d[m_s \times a_s]) \\ &+ \frac{\alpha}{2} \int_{\mathcal{M} \times \mathcal{A}} \{-v(q(z_b, z)) + W(m + d(z_b, z) + a + \tau; G, \theta)\} F(d[m_b \times a_b]) \\ &+ (1 - \alpha) W(m + a + \tau; G, \theta), \end{aligned} \tag{9}$$

where, as noted before, $G(m, a) = \Gamma_G(F(m, a), \theta)$. Here, the first term is the expected value of being matched as a buyer (taking into account that there are several "types" of sellers they can meet with), the second term is the expected value of being matched as a seller, and the last term is the value of not being matched or being matched in a no-coincidence meeting.

2.3 Centralized Market

Recall that agents enter the CM with a certain level of wealth, a variable that includes nominal assets at the end of the previous DM (i.e., net of decentralized trading payments) and government transfers. Thus, an agent with wealth x at the beginning of the CM has lifetime utility given by:

$$W(x; G, \theta) = \max_{c,h,m',a'} \left\{ U(c,h) + \beta V\left(m', a'; F', \theta'\right) \right\} \tag{10}$$

subject to

$$c = h + \phi_m(G,\theta)\, x - \phi_m(G,\theta)\left[m' + \phi_a(G,\theta)\, a'\right]\left(1+\mu^M\right) \tag{11}$$

$$F'\left(m', a'\right) = \Gamma_F\left(G(m,a), \theta\right). \tag{12}$$

2.4 Equilibrium

A monetary equilibrium is a set of functions for value $\{V(m,a;F,\theta), W(x;G,\theta)\}$, allocations $\left\{c(x;G,\theta), h(x;G,\theta), m'(x;G,\theta), a'(x;G,\theta), q(z_b,z_s)\right\}$, prices $\left\{d(z_b,z_s), \phi_m(G), \phi_a(G)\right\}$, and distributions $\left\{F(m,a), G(m,a)\right\}$ such that, given the policy θ:

1. Values $V(m,a;F,\theta)$ and $W(x;G,\theta)$ and decision rules $c(x;G,\theta)$, $h(x;G,\theta)$, $m'(x;G,\theta)$, $a'(x;G,\theta)$ satisfy the definitions above, for any given $\{q, d, \phi_m, \phi_a\}$ and $\{F(m,a), G(m,a)\}$.

2. Terms of trade $\{q(z_b,z_s), d(z_b,z_s)\}$ in the DM solve the problem in (7) given $V(m,a;F,\theta)$ and $W(x;G,\theta)$.

3. There is a monetary equilibrium: $\phi_m > 0$.

4. The government has a balanced budget. This is, Equation (2) holds.

5. The money and bond markets clear, i.e., Equations (5) and (6) are satisfied.[3]

6. The law of motions for $F(m,a)$ and $G(m,a)$ are given by $\Gamma_F(\cdot)$ and $\Gamma_G(\cdot)$, respectively. These maps are consistent with the initial conditions and the evolution of

[3] Eq. (4) shows that there is a one-to-one mapping between the price of nominal bonds, ϕ_a, and the nominal interest rate, i. This, together with the market clearing condition in the bonds market, implies that there is a direct relation between i and the relative supply of bonds, K. Thus, in the stationary equilibrium, for any policy $\theta = (K, \mu)$, there is an equivalent policy (i, μ).

money holdings implied by DM and CM trade. This is,

$$\begin{aligned} G(m,a) = \Gamma_G(F,\theta) = &\frac{\alpha}{2}\int_{\{z_s\in Z\}}\int_{\{(m-d(z,z_s)+\tau,a)\in Z\}} F([dm\times da])\,F([dm_s\times da_s]) \\ &+\frac{\alpha}{2}\int_{\{z_b\in Z\}}\int_{\{(m+d(z_b,z)+\tau,a)\in Z\}} F([dm\times da])\,F([dm_b\times da_b]) \\ &+(1-\alpha)\int_{\{(m+\tau,a)\in Z\}} F([dm\times da]) \end{aligned} \tag{13}$$

and

$$F'(m,a) = \Gamma_F(G,\theta) = \int_{\{(m'(x),a'(x))\in Z\}} G([dm\times da]), \tag{14}$$

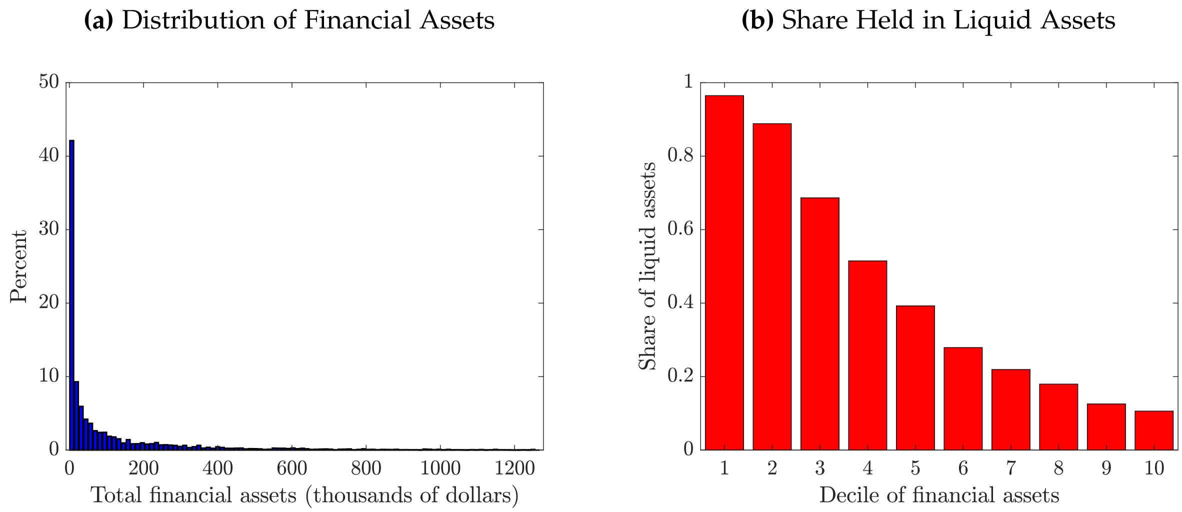
where, as before, $z_i = (m_i, a_i)$ for $i = b, s$ and $Z = \mathcal{M}\times\mathcal{A}$.

3 Data and Calibration

The quantitative results of the model regarding the effectiveness of OMOs depend on how closely the model resembles the distribution of financial assets in the data. In the same spirit, it is necessary to match agents' portfolio composition along the distribution of assets. Figure 2(a) presents the distribution of financial assets computed using data from the SCF for 2007. This distribution exhibits a long right tail, and there is a large concentration of households at lower levels of financial assets. Consequently, the Gini coefficient for this distribution is 0.86. Figure 2(b) shows the share of financial assets held in liquid assets by decile. The most asset-poor households (decile 1) tend to hold close to 95 percent of their financial assets in liquid instruments. In comparison, the most asset-rich households only hold slightly more than 10 percent in liquid assets on average.

The model is parametrized such that one model period is equivalent to one year. Hence, the discount factor is set to $\beta = 0.96$. Aggregate equilibrium objects target averages for the period 1984Q1–2007Q4. The money growth rate is consistent with an inflation rate, in the stationary equilibrium, equal to the average of 3.1 percent in the data. To measure money in the data, I follow Alvarez, Atkeson and Edmond (2009), where money is computed as the sum of currency, demand deposits, saving deposits, and time deposits. For the period of study, the ratio of the total federal debt to money is 1.28. In the model, this is equivalent to K. I use this as a target to calibrate the probability of being matched in a DM meeting, α. Regarding interest rates for government debt, I

Figure 2: Distribution and Composition of Financial Assets in the SCF 2007



Notes: Data from the Survey of Consumer Finances for 2007. Liquid assets computed as the sum of money market, checking, savings and call accounts, and prepaid cards. Financial assets include all liquid assets plus certificates of deposit, directly held pooled investment funds, saving bonds, directly held stocks and bonds, cash value of whole life insurance, other managed assets, quasi-liquid retirement accounts, and other financial assets. The share of liquid assets is computed as the ratio of liquid assets to total financial assets.

use the secondary market rate of the 3-month Treasury bills. The average of this nominal rate for the same period is 4.9 percent.

The parameters for the scale and curvature of the cost of working in the DM, B and ν, are set to match the average markup and the velocity of money. The model is able to replicate a velocity of money close to the one observed in the data for the sample period: 2.28 in the model versus 2.25 in the data. Moreover, the average markup[4] generated in DM meetings (32 percent), is consistent with the 30 percent average markup reported in Faig and Jerez (2005), while the semi-elasticity of money demand with respect to the nominal interest rate is -0.69, close in range to the one documented in Lucas (2000) and Berentsen, Menzio and Wright (2011). Table A1 in Appendix A.2 presents the local elasticities of some of the model's moments with respect to the parameters of this baseline calibration.

[4]The markup is computed as the ratio of the price and the marginal cost of production of the quantities agreed in a given meeting. The average markup weighs the set of possible individual markups for the occurrence probabilities of those meetings.

Table 1: Parameter Values

Parameter		Value
α	Probability of meeting	0.90
β	Discount factor	0.96
$\overline{\mu}$	Growth rate of money and bonds	0.031
Decentralized market		
η	Curvature of utility of consumption	0.99
b	Scale parameter in $u(c_d)$	0.0001
B	Scale of cost of working	0.4
ν	Curvature of cost of working	1.15
Centralized market		
κ	Scale of disutility of labor	2.0
χ	Inverse of Frisch elasticity	2.0

In terms of functional forms and model parameters, I follow Lagos and Wright (2005) for the utility and cost functions in the DM:

$$u(c_d) = \frac{1}{1-\eta}\left[(c_d + b)^{1-\eta} - b^{1-\eta}\right] \tag{15}$$

$$v(h_d) = Bh_d^{\nu} \tag{16}$$

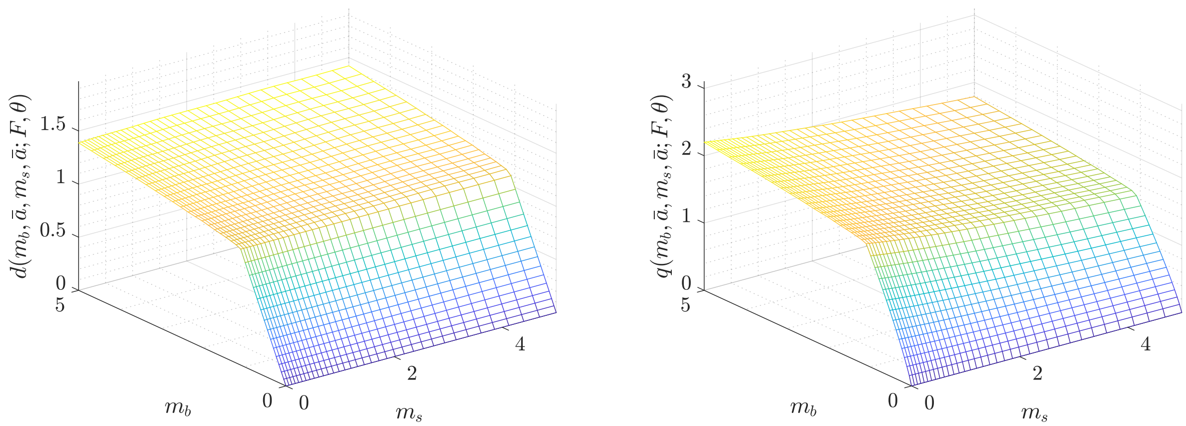
with $\eta > 0$, $B > 0$, $\nu > 0$, and $b \approx 0$. However, for the CM, I drop the quasi-linearity assumption. In particular, the utility function in the CM is concave in consumption, has a constant Frisch elasticity of labor supply, and is given by:

$$U(c,h) = \log c - \kappa \frac{h^{1+\chi}}{1+\chi}, \tag{17}$$

where χ is the inverse of the Frisch elasticity of labor supply and $\kappa > 0$ a scale parameter. The remaining parameters are chosen to illustrate the core mechanisms of the model. The parameter that governs the curvature of the utility function in the DM is set close to 1, so the utility function is close to being logarithmic. The remaining parameters are summarized in Table 1.

Figure 3: Terms of Trade: Monetary Payments and Quantities

(a) Payments **(b)** Quantities



Notes: For expositional purposes, the figures only present terms of trade as a function of (m_b, m_s) for meetings between both buyers and sellers with average bond holdings, $\bar{a}$. See Figure A1 in Appendix A.2 for more figures emphasizing the role of bond holdings in determining terms of trade in the decentralized market.

4 Properties of Stationary Equilibrium

Before examining the role of OMOs, I first characterize the stationary equilibrium by showing how the idiosyncratic liquidity risk that arises from decentralized trading shapes terms of trade and the distribution of assets.[5]

Figure 3 shows the terms of trade for quantities and monetary payments exchanged in meetings between buyers and sellers with average bond holdings. The figure depicts terms of trade as a function of the buyer's money holdings, m_b, and the seller's, m_s. In any possible meeting, the quantities exchanged in the DM are an increasing function of m_b and depend negatively on m_s. In addition, monetary payments increase in both m_b and m_s. Intuitively, the wealthier the seller is, the higher is his opportunity cost of working. Hence, the seller requires more money in exchange for a lower level of output in order to remain indifferent about trading. On the other hand, the buyer's willingness to pay increases with her money holdings as long as it guarantees her higher consumption, i.e., the opportunity cost of spending an extra dollar in DM trading decreases with m_s.

[5]For all these numerical experiments, I use the solution method outlined in Appendix A.1.

There is a similar outcome in terms of bond holdings. Wealthier buyers are willing to pay more, and sellers with higher bond holdings have a higher opportunity cost.[6]

Note that, as is standard in search-theoretic models of money, trading frictions in decentralized trading generate inefficient allocations. The efficient level for quantities exchanged, q^*, solves $u'(q^*) = v'(q^*)$. However, differences in total wealth and not only differences in money holdings are the ones giving rise to inefficient outcomes.[7] For the benchmark calibration of the model, $q^* = 1.98$. Thus, as shown in Figure 3(b) for the case of agents with average bond holdings, meetings in which a wealthy buyer trades with a poor seller result in an inefficiently high level of production. Conversely, meetings with poor buyers tend to generate an inefficiently low level of trade. This means that a higher dispersion in money and bond holdings generates a more inefficient equilibrium.

The various possible outcomes in decentralized trading lead to buyers and sellers entering the next CM with different asset holdings and portfolio compositions. As opposed to models where preferences are quasi-linear, wealth effects play a significant role in determining agents' decisions in the CM, and not all agents choose to accumulate the same amount of money or bonds (see Figures A2 and A3 in Appendix A.2). As a consequence, the heterogeneity intrinsically generated by decentralized trading is persistent across periods.

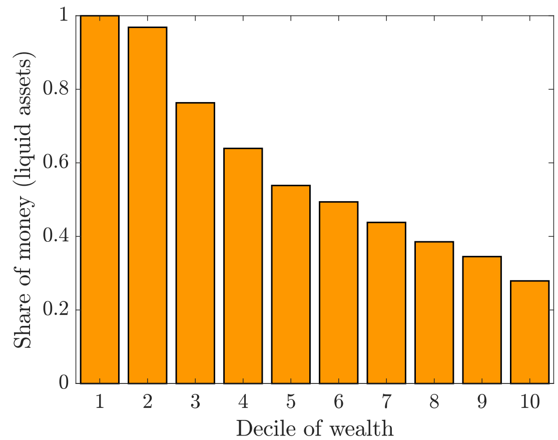
Interestingly, this mechanism generates a composition of portfolios consistent with the data presented before in Figure 2(b). Figure 4 presents the share of total assets held as money, along the different levels of wealth (total assets or cash). This figure shows how asset-poor agents prioritize the accumulation of money over bonds as they prepare to enter a new round of decentralized trading (see Figure A2). On the other hand, asset-rich agents can self-insure more easily by purchasing a higher share of bonds.

Figures 5(a) and 5(b) depict the distribution of agents over money and government bonds at the beginning of the DM and the CM. The distribution at the beginning of the DM is relatively concentrated in terms of money holdings but disperse in bonds. In contrast, the distribution entering the CM spreads agents out with respect to their money holdings depending on if they are buyers, sellers, or unmatched in the previous DM. Why is there such a drastic change in the dispersion of money but not in the disper-

[6] See Figure A1 in Appendix A.2 for a summary of the quantities traded, monetary payments, and prices per unit for more types of meetings in which bond-rich and bond-poor agents meet.

[7] Let γ be the multiplier on the buyer's liquidity constraint and define wealth as $x_i = m_i + a_i + \tau$ for $i = b, s$. The optimality conditions for the bargaining problem imply $u'(q)/v'(q) = [W'(x_b - d) + \gamma]/W'(x_s + d)$. Given that W is monotonically increasing and concave, if $x_b < x_s$, then $W'(x_b - d) > W'(x_s + d)$, which implies that $u'(q) > v'(q)$ and $q < q^*$.

Figure 4: Share of Liquid Assets in Portfolio



Notes: Shares of liquid assets in agents' portfolios by decile in the model's stationary equilibrium. The share of liquid assets is computed as $m' / (m' + a')$.

Figure 5: Distribution of Assets

(a) At the Beginning of DM

(b) At the Beginning of CM

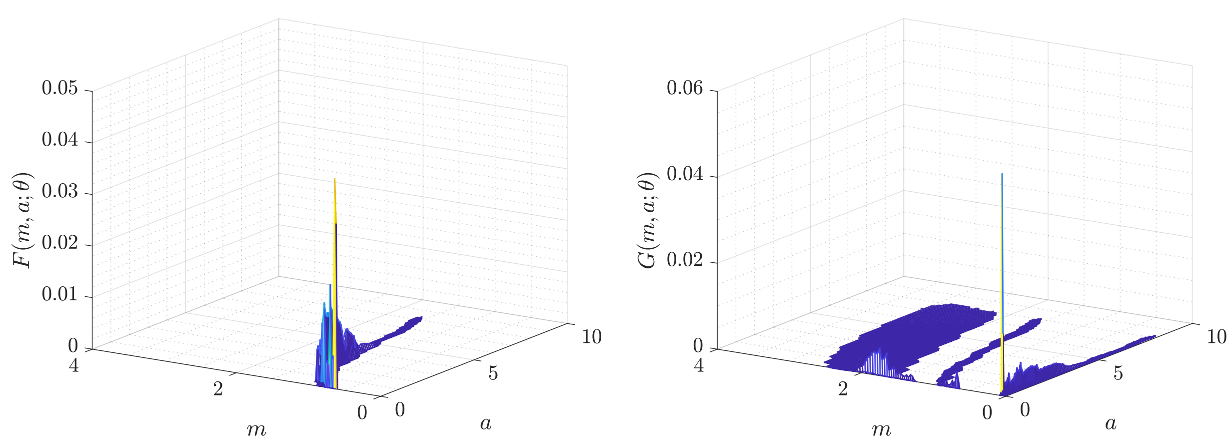
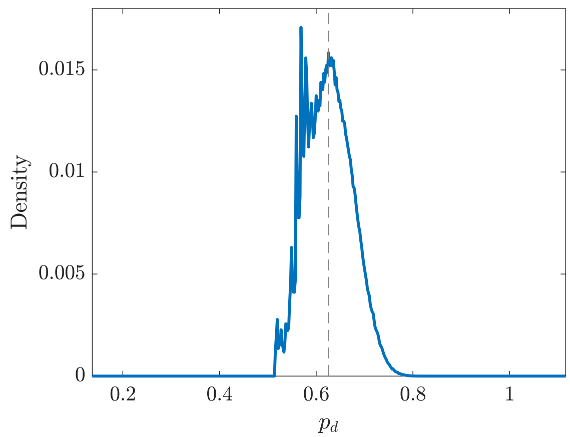


Figure 6: Stationary Distribution of Prices

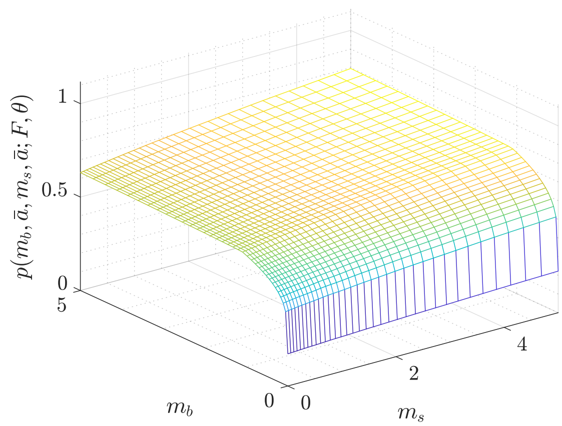


Notes: For a particular meeting with buyer's and seller's portfolios $z_b = (m_b, a_b)$ and $z_s = (m_s, a_s)$, respectively, prices are defined as $p(z_b, z_s) = d(z_b, z_s) / q(z_b, z_s)$. The distribution of prices accounts for all possible meetings between buyers and sellers given the stationary distribution of agents over money and bond holdings.

sion of bonds? We are observing risk-averse agents using bonds to partially self-insure against the idiosyncratic liquidity shocks they may experience in the DM. However, the reshuffling of a poor agent's portfolio is not immediate because the marginal cost of working is increasing in the hours worked. As a result, agents adjust their money balances very promptly because of the risk of being matched as buyers in the next period. Moreover, agents accumulate bonds at a slower pace after having reached some level of money that is enough to face a liquidity shock in the next DM. This pattern is clear when we observe the decision rules for $m' = g_m(x; G, \theta)$ and $a' = g_a(x; G, \theta)$ in Figure A2 in Appendix A.2.

This heterogeneity across agents' portfolios when entering the DM generates a fully-fledged distribution of prices, where the price per unit is defined as $p_d = d/q$. Figure 6 shows such distribution of prices. Given the relatively high concentration of agents around a particular value of m, there is a large fraction of meetings with a price around the mean price (0.63). However, the variance in bond holdings is sufficient to generate a standard deviation of 0.048 in the observed prices. As with the terms of trade for quantities and payments, in Figure 7, I compute prices per unit as a function of both agents' money holdings while fixing their stock of bonds at the average level. One of the most relevant variables when determining prices is how costly it is for the seller to produce. A seller with high levels of money and bond holdings requires exchanging at

Figure 7: Terms of Trade: Prices



Notes: Possible outcomes for meetings between buyers and sellers with average bond holdings, $\overline{a}$. Prices are defined as $p(z_b, z_s) = d(z_b, z_s)/q(z_b, z_s)$, where buyers' and sellers' portfolios are given by $z_b = (m_b, \overline{a})$ and $z_s = (m_s, \overline{a})$, respectively.

higher prices in decentralized trading because of his participation constraint. The fact that, in equilibrium, the model exhibits a large dispersion in asset holdings explains the variance in the distribution of prices. Hence, as long as the distribution of money and bonds has a sluggish response to economy-wide shocks (e.g., OMOs), we can expect that the distribution of prices will react accordingly.

5 The Effects of Open Market Operations

Now I turn to investigate the role of OMOs. First, I examine the stationary equilibrium of the model under different targeted nominal interest rates. As discussed before, each level for the targeted rate requires the consolidated fiscal and monetary authority to adjust the public provision of assets. I assume that the government implements a particular level of $K = A/M$ by conducting an OMO and then study the long-run properties of such a policy. Second, I assess how these stationary equilibria change for different trend inflation values, $\overline{\mu}$.

5.1 Open Market Operations in the Long Run

Table 2 presents some macroeconomic aggregates for three different stationary equilibria: the baseline calibration in Section 4, where the targeted rate is 4.9 percent, and two additional ones with targeted rates 1 percentage point (p.p.) above and below the baseline level. As suggested by the theoretical models of Kocherlakota (2003) and Shi (2008), implementing higher interest rates on public debt requires an increase in the relative supply of government bonds. This is simply a natural response from the demand for bonds: for a given level of the trend inflation, an increased supply of bonds pushes their price down, which in turn increases the nominal interest rate. As a consequence, higher nominal rates are associated with increased overall liquidity in the economy.[8] Meanwhile, when interest rates are high, the relative scarcity of money translates into an increased price of money.

The increased overall liquidity is associated with more economic activity. In particular, while there is a slight change in the CM's level of activity, output in the DM falls 0.9 percent when the interest rate goes from 4.9 to 3.9 percent and increases 0.6 percent when we compare the benchmark scenario with the case where the interest rate is 5.9 percent. The relatively small increase in the CM's production for higher levels of aggregate liquidity comes from the fact that the wealth effect associated with a higher disposable income is partially canceled out—in the aggregate level—by the substitution effect originating from the increased prices of money and nominal bonds.

In contrast, aggregate output from DM meetings increases with increased total liquidity levels because wealth becomes more evenly distributed. While the dispersion in bond holdings increases, this occurs in a smaller proportion than the rise in the relative supply of bonds. Likewise, given the increased supply of assets that let agents hedge their idiosyncratic risk, and the subsequent drop in the price of public debt, it becomes more attractive for agents to have a portfolio with lower participation of money. This explains the reduced dispersion in money holdings in the DM, as well as in the CM. This contraction in both dimensions of wealth reduces the inefficiencies in DM trading discussed earlier so that any pair of agents meeting in the DM look more alike and exchange more DM goods at lower and less disperse prices.

These results also highlight the presence of nonlinearities. Compared with the baseline scenario, targeting an interest rate of 1 p.p. higher or lower requires a change in the relative supply of bonds of 22 and -20.9 percent, respectively. However, although these two changes are relatively similar, total output contracts 0.31 percent in the economy with reduced liquidity, significantly more (in absolute value) than the increase of 0.21 percent for the case with a higher supply of bonds. Most of this difference is explained by the changes in DM activity. As agents have fewer chances of self-insuring, there is

[8] Total liquidity is defined as the sum of the nominal value of the aggregate supply of money and government bonds. This is $1 + \phi_a K$.

Table 2: Stationary Equilibrium for Different Nominal Interest Rates

Nominal interest rate	-1**p.p.**	**Baseline** 4.9%	$+1$**p.p.**
Ratio of bonds to money	0.915	1.157	1.411
Total output	1.310	1.314	1.317
Output in DM	0.512	0.516	0.520
Output in CM	0.798	0.798	0.797
Std. dev. of money in DM	0.193	0.171	0.158
Std. dev. of money in CM	0.952	0.960	0.960
Std. dev. of bonds	0.866	1.015	1.167
Average DM price	0.625	0.626	0.619
Std. dev. of DM price	0.049	0.048	0.046

Notes: Stationary equilibria for three different targeted nominal interest rates under the baseline calibration. The baseline scenario targets a 4.9% rate, while the two additional scenarios target a 1 p.p. change relative to the baseline rate (3.9% and 5.9%).

more risk associated with bilateral meetings and an increased incidence of inefficiencies in DM trading.

Despite the relatively small changes in the CM's activity at the aggregate level, there are significant differences between these stationary equilibria along the distribution of agents. Table 3 shows the long-run changes in CM consumption, labor, and asset holdings, as well as DM consumption, for different groups of agents divided by wealth, after an OMO that targets a higher interest rate.[9]

The poorest agents gain the most when the economy moves to a scenario of increased liquidity with higher interest rates. Agents below the 25th percentile consume more and work less in the CM. Similarly, agents in the two lowest quartiles increase their money holdings, while agents between the 10th and 50th percentiles also experience higher growth in their bond holdings than those at the top half of the distribution.[10] Note that it is still optimal for agents in the lowest decile to have a portfolio with no bonds in it.

The reduced concentration of assets also benefits the agents at the lowest end of the wealth distribution in their DM meetings. As shown in Table 2, agents trade more at lower and less dispersed prices. This implies that, primarily, agents in the lowest half of

[9]See Table A2 in Appendix A.2 for the long-run results of an OMO that lowers the interest rate in 1 p.p.

[10]The higher supply of government bonds is distributed across almost the entire population. All agents, except the ones in the first decile, increase their bond holdings.

Table 3: Long-Run Distributional Effects of an Open Market Operation from $i = 4.9\%$ to $i = 5.9\%$

Percentile	≤ 10	**10-25**	**25-50**	**50-75**	**75-90**	≥ 90
Consumption (CM)	1.33	2.25	0.06	-0.47	-0.99	-0.77
Labor	-0.62	-1.12	-0.03	0.22	0.51	0.39
Money holdings	2.37	2.40	1.64	-0.75	-2.52	-0.75
Bond holdings	-	117.75	31.42	22.66	19.75	16.03
Consumption (DM)	4.02	4.36	2.19	-0.35	-1.76	-0.25

Notes: Percent changes by percentile groups across stationary equilibria with different targeted nominal interest rates. Changes are with respect to the baseline economy with a 4.9% nominal interest rate. Agents in the lowest decile have zero bond holdings in both stationary equilibria.

the distribution obtain better terms of trade and are able to consume more when they are buyers in bilateral meetings.

5.2 The Role of Trend Inflation

The level of trend inflation also matters when determining the long-run effects of OMOs at the aggregate level, as well as at the distributional one. Table 4 compares key macroeconomic variables for different stationary equilibria with interest rates of 4.9 and 5.9 percent and trend inflation levels of 2.1 and 3.1 percent.[11] When money and bonds grow at a faster pace, assets lose value more rapidly, and agents trade at higher prices in their DM meetings. This also implies that agents are in higher need of self-insuring against the idiosyncratic liquidity shocks they may experience in DM trading. Thus, the increased demand for bonds lowers the supply needed to implement a given nominal interest rate. Therefore, an economy with higher trend inflation has less liquidity, lower output, and a higher concentration of wealth.

Trend inflation also has an impact on the long-run redistributive effects of an OMO. Figure 8 shows how much CM consumption changes by wealth percentile groups. For most agents, consumption does not change significantly. However, having the increased supply of liquidity needed to implement a higher interest rate has a large positive impact on the poorest agents. As shown in Table 4, the supply of bonds needs to increase relatively more for the economy with lower trend inflation. This tends to benefit more

[11] Recall that the baseline economy is calibrated with a level of trend inflation of 3.1 percent and a nominal interest rate of 4.9 percent. Hence, column three of Table 4 coincides with column two of Table 2.

Table 4: Long-Run Effects of Open Market Operations for Different Levels of Trend Inflation

	$\overline{\mu} = 2.1\%$		$\overline{\mu} = 3.1\%$	
Nominal interest rate	4.9%	5.9%	4.9%	5.9%
Ratio of bonds to money	1.887	2.855	1.157	1.411
Total output	1.331	1.343	1.314	1.317
Output in DM	0.534	0.543	0.516	0.520
Output in CM	0.797	0.800	0.798	0.797
Std. dev. of money in DM	0.144	0.118	0.171	0.158
Std. dev. of money in CM	0.958	0.944	0.960	0.960
Std. dev. of bonds	1.414	1.881	1.015	1.167
Average DM price	0.599	0.589	0.626	0.619
Std. dev. of DM price	0.044	0.043	0.048	0.046

Notes: Stationary equilibria for two different levels of trend inflation, $\overline{\mu}$, and two different targeted nominal interest rates.

the most liquidity-constrained agents, as the concentration of assets becomes more even and terms of trade move closer to the efficient allocation.[12]

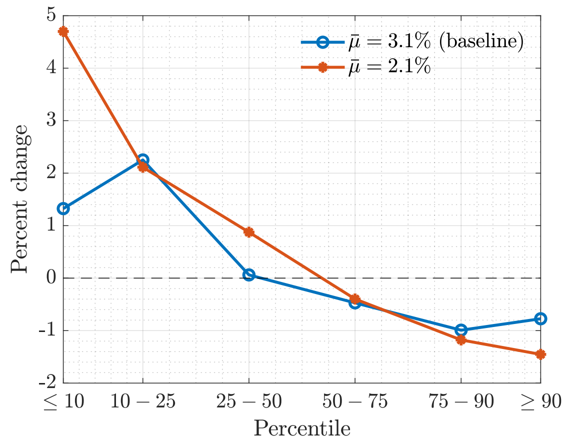
5.3 Optimal Monetary Policy

OMOs that increase the relative supply of bonds also increase total output in a zero-inflation economy, as well as in economies with large levels of trend inflation. Table 5 presents the long-run effects of OMOs that target a 1 p.p. increase of the nominal interest rate when trend inflation is 0 percent and 10 percent. For both cases, in the initial stationary equilibrium, the ratio of bonds to money is set as close as possible to the one in the baseline calibration, $K = 1.16$, and the nominal interest rate is then determined endogenously. As before, implementing an OMO that increases the nominal interest rate in 1 p.p. requires solving for the level of the supply of bonds that is consistent with that new rate.

As shown in the previous section, the more rapid the growth of nominal assets, the less attractive they are as they lose value faster. Thus, higher inflation levels are associated with lower output, higher DM prices, and an increased concentration of wealth.

[12]There is a special case when the real return on bonds is negative, i.e., when $\overline{\mu} > i$. In this case, agents have no incentives to have bonds in their portfolios and hence only demand money, as it still gives them the liquidity yield associated with decentralized trading. When this happens, the economy collapses to one in which money is the only asset and $K = A/M = 0$. See Appendix B for an economy for which this is the case.

Figure 8: Long-Run Distributional Effects on Centralized Market Consumption of an Open Market Operation and Trend Inflation



Notes: Percent changes by percentile groups across stationary equilibria with different targeted nominal interest rates, under two different levels of trend inflation, $\overline{\mu}$. Changes are from an economy with a 4.9% nominal interest rate to one with a 5.9% rate.

It also holds that OMOs that target a higher interest rate require increasing the relative supply of nominal bonds. Such an increase is larger for lower levels of trend inflation. As reported in Table 5, the required increase is five times higher when inflation is 0 percent than when it is 10 percent. Likewise, these OMOs increase output, reduce the concentration of wealth, and lower DM prices in the long run.

Importantly, these results are also related to the optimal conduct of monetary policy in the stationary equilibrium. Columns 1 and 3 of Table 5 and Column 3 of Table 4 show the stationary equilibrium for trend inflation equal to 0 percent, 10 percent, and 3.1 percent, respectively, when the relative supply of bonds is set to its calibrated value. When trend inflation is zero, and the relative supply of bonds is set to such value, the implied equilibrium nominal interest rate is roughly zero.[13] For the calibrated level of K, the output-maximizing monetary policy is the Friedman rule: $(i, \overline{\mu}) = (0, 0)$. This policy is also welfare-maximizing, as presented in Table 6 in the next section.

However, and in contrast to Lagos and Wright (2005), implementing the Friedman rule in this model economy does not imply eliminating the inefficiencies that arise in decentralized trading. These inefficiencies are always present at any level of inflation or

[13]For other levels of K, having zero trend inflation does not necessarily imply a zero nominal interest rate. However, for any level of $K > 0$ that is consistent with the existence of a monetary equilibrium, there is a level of $\overline{\mu}$ such that the nominal rate is zero.

Table 5: Long-Run Effects of Open Market Operations for Low and High Levels of Trend Inflation

	$\overline{\mu} = 0\%$		$\overline{\mu} = 10\%$	
Nominal interest rate	0.02%	1.02%	19.9%	20.9%
Ratio of bonds to money	1.170	2.298	1.156	1.380
Total output	1.352	1.357	1.253	1.254
Output in DM	0.555	0.558	0.457	0.457
Output in CM	0.797	0.799	0.797	0.797
Std. dev. of money in DM	0.237	0.133	0.153	0.138
Std. dev. of money in CM	0.802	0.943	0.963	0.956
Std. dev. of bonds	0.992	1.638	1.011	1.133
Average DM price	0.475	0.573	0.699	0.690
Std. dev. of DM price	0.037	0.044	0.047	0.045

Notes: For both levels of $\overline{\mu}$, the initial stationary equilibrium is computed while imposing that the ratio of bonds to money is as close as possible to the one in the baseline economy, $K = 1.16$. The second stationary equilibrium solves for the required level of K that implements an increase of 1 p.p. in the nominal interest rate.

nominal interest rate. This happens because as long two agents with different levels of wealth are matched in the DM, the quantity they trade is below the efficient level, i.e., $q(z, \tilde{z}) < q^*$ for $z \neq \tilde{z}$ as discussed in Footnote (7). Likewise, the presence of wealth effects in the CM, and the fact that not all agents rebalance their portfolios immediately after trading in the DM, guarantees that there will be persistent heterogeneity. Therefore, in this model the Friedman rule is not optimal in the same sense as in Lagos and Wright (2005) where it implies that $q(z, \tilde{z}) = q^*$.

6 Welfare Cost of Inflation

A traditional question in the monetary economics literature has to do with the welfare cost of inflation. The welfare cost is measured in terms of consumption compensation. It considers how much consumption agents would sacrifice for not moving from a zero inflation equilibrium to one with positive inflation, usually 10 percent. Cooley and Hansen (1989) estimate this cost to be 0.4 percent of total output, while Lucas (2000) reports that it is slightly less than 1 percent, and Lagos and Wright (2005) suggest that it is 1.4 percent of consumption for their benchmark model. These papers, among many others, have traditionally relied on environments without persistent heterogeneity. In this type of setting, one of the major determinants of the cost of inflation is the inflation tax, i.e., the reduced incentive to hold cash and substitute away from activities that require it when inflation is higher.

In the model presented above, there are two conflicting forces. On the one hand, with higher inflation, agents holding more nominal assets experience a faster contraction in the real value of their portfolio. This means that the incidence of the inflation tax is not evenly distributed over the population when we have a nondegenerate distribution of money and nominal bonds. Moreover, part of the seigniorage money and new debt issuance are used to cover interest payments on government debt. Given that agents tend to hold just the money they need for transactional purposes, the concentration in bond holdings is key in determining which agents are receiving those interest payments. On the other hand, higher inflation rates are associated with larger lump-sum transfers. Hence, agents with low money holdings will then be benefited as their marginal valuation of one extra unit of money is higher than the one for rich agents. These mechanisms highlight the importance of distributional effects when measuring the cost of inflation.

In an environment similar to the one presented here, with persistent heterogeneity but with only one asset, Chiu and Molico (2011) estimate that the cost of inflation is even less than what Lucas (2000) documents because of the redistribution of resources towards the poorest agents. However, in the presence of only one asset, all seigniorage money is evenly distributed across agents. I follow a similar approach to theirs when computing the cost of inflation. First, note that the agents' average expected lifetime value when the steady-state inflation rate is μ is given by:

$$\mathfrak{U}(\mu) = \frac{1}{(1-\beta)}\left[\frac{\alpha}{2}\int_Z\int_Z \{u(q(z,\tilde{z};\mu)) - v(q(\tilde{z},z;\mu))\} F([dm\times da];\mu) F([d\tilde{m}\times d\tilde{a}];\mu) \right. \\ \left. + \int_Z U(c(z;\mu), h(z;\mu)) G([dm\times da];\mu)\right], \tag{18}$$

where $Z = \mathcal{M}\times\mathcal{A}$. Then, the welfare cost of having an inflation of μ with respect to 0 is $\lambda(\mu) - 1$. This value captures how much of their consumption agents would be willing to sacrifice to not have an inflation rate equal to μ. We can compute $\lambda(\mu)$ by solving:

$$\mathfrak{U}(\mu) = \frac{1}{(1-\beta)}\left[\frac{\alpha}{2}\int_Z\int_Z \{u(q(z,\tilde{z};0)\lambda_0(\mu)) - v(q(\tilde{z},z;0))\} F([dm\times da];0) F([d\tilde{m}\times d\tilde{a}];0) \right. \\ \left. + \int_Z U(c(z;0)\lambda_0(\mu), h(z;0)) G([dm\times da];0)\right]. \tag{19}$$

Table 6 presents the computed values for $\lambda(\mu)$ or different values of the steady-state inflation rate, μ. For the case of 10 percent inflation, for example, the welfare cost is 1.67 percent, slightly higher than that in the comparable calibration in Lagos and Wright (2005), and almost twice as much as in Chiu and Molico (2011). As discussed before, with higher inflation, agents are less interested in holding money balances, so the dispersion at the beginning of the DM falls. In contrast, there is an increased demand for bonds that pushes the dispersion in bond holdings upward and raises wealth inequality. As a result, DM prices increase in variance and in level.

To understand these results, it is helpful to measure how much of the welfare changes can be attributed to changes in the distribution of agents, and to how much they adjust their behavior. For low levels of inflation, changes in the distribution alone end up contributing to an increase in the welfare cost of inflation, as the dispersion in bond holdings and wealth increases, even if we maintain the same terms of trade and decision rules from the zero-inflation equilibrium. However, as inflation rises, the reduced concentration in money holdings dominates the increase in overall inequality and therefore reduces the cost of inflation. On the other hand, when focusing on the decision rules for the cases in which $\overline{\mu} > 0$, agents value less their nominal asset holdings with higher levels of trend inflation. This implies that they are willing to trade bilaterally at higher prices while wanting to accumulate more bonds that give them some positive real return. All these shifts in agents' behavior increase the welfare cost of inflation.

7 Concluding Remarks

In this paper, I study the long-run effects of OMOs while accounting for the interactions between aggregate variables and the distribution of assets across households. To do this, I develop a search-theoretic model of money where a consolidated fiscal and monetary authority can use OMOs to manage the public provision of liquidity. The model generates a nondegenerate distribution of money and bonds as agents do not adjust their asset holdings immediately after an idiosyncratic liquidity shock. Notably, the portfolio composition for agents along the distribution of wealth reproduces the one observed in the data.

My results show that a permanent increase in the relative supply of government bonds is associated with higher nominal interest rates and more economic activity. The increased supply of bonds reduces the concentration in asset holdings and total wealth while also expanding the chances for agents to self-insure against idiosyncratic liquidity

Table 6: Welfare Cost of Inflation Relative to 0% Trend Inflation

Inflation rate, $\overline{\mu}$	**0%**	**1%**	**2%**	**5%**	**10%**
Welfare cost (%)	–	0.121	0.355	0.758	1.666
Welfare change from distributions	–	-0.064	-0.024	0.023	0.067
Welfare change from decision rules	–	-0.213	-0.442	-0.729	-1.171
Average DM price	0.475	0.563	0.608	0.646	0.699
Std. dev. of DM price	0.037	0.044	0.047	0.048	0.047
Std. dev. of money in DM	0.237	0.189	0.178	0.168	0.153
Std. dev. of money in CM	0.802	0.904	0.953	0.965	0.963
Std. dev. of bonds	0.992	1.007	1.021	1.015	1.011

Notes: The welfare cost of inflation is measured as the percentage of consumption, in both the centralized and decentralized markets, that agents are willing to sacrifice to not have an inflation rate of $\overline{\mu}$, relative to a zero-inflation equilibrium. The welfare change explained by changes in the distributions measures changes in agents' average utility, while leaving the decision rules and DM terms of trade constant as in the equilibrium with inflation $\overline{\mu} = 0$. The welfare change explained by changes in decision rules does the same but keeps the distributions F and G as in the equilibrium with zero inflation. Each column is obtained by computing the stationary equilibrium for that particular level of the inflation rate $\overline{\mu}$, while imposing the ratio of bonds to money from the baseline economy, $K = 1.16$.

shocks. Furthermore, as agents become more alike in terms of asset holdings, the incidence of frictions in bilateral trading shrinks so that agents, on average, trade at a level closer to the efficient one and at lower and less dispersed prices. As a result, total output increases, and the poorest and more liquidity-constrained agents gain the most, as some resources get redistributed out of the wealthiest agents in the economy.

I also find asymmetries in the response of the economy to OMOs. When targeting interest rate hikes, the size of the required OMO is larger than when implementing an equivalent interest rate contraction. Moreover, this asymmetry also affects total output. By cutting the public provision of liquidity, agents become less able to self-insure. This, in turn, exacerbates the negative consequences of liquidity shocks while increasing the probability of having DM meetings with trade further away from the efficient level. Such an effect is stronger when the supply of liquidity falls, as more agents are pushed to be liquidity constrained.

Finally, the presence of heterogeneity across agents and the existence of a fully-fledged distribution of prices suggest that a one-time permanent OMO has enough room to generate short-run non-neutralities. In particular, since agents cannot rebalance their portfolio instantaneously after a liquidity shock, the distribution of assets and the distribution of prices would slowly respond to the increased liquidity. This extension is worth considering in future research.

References

Alvarez, Fernando, Andrew Atkeson, and Chris Edmond, "Sluggish Responses of Prices and Inflation to Monetary Shocks in an Inventory Model of Money Demand," *The Quarterly Journal of Economics*, 2009, *124* (3), 911–967.

Araujo, Luis and Leo Ferraris, "Money, Bonds, and the Liquidity Trap," *Journal of Money, Credit and Banking*, 2020, *52* (7), 1853–1867.

Berentsen, Aleksander, Guido Menzio, and Randall Wright, "Inflation and Unemployment in the Long Run," *American Economic Review*, 2011, *101*, 371–398.

Carli, Francesco and Pedro Gomis-Porqueras, "Real Consequences of Open Market Operations: The Role of Limited Commitment," *European Economic Review*, 2021, *132*, 103639.

Chiu, Jonathan and Miguel Molico, "Liquidity, Redistribution, and the Welfare Cost of Inflation," *Journal of Monetary Economics*, 2010, *57* (4), 428–438.

_ and _ , "Uncertainty, Inflation, and Welfare," *Journal of Money, Credit and Banking*, 2011, *43*, 487–512.

_ and _ , "Short-Run Dynamics in a Search-Theoretic Model of Monetary Exchange," *Review of Economic Dynamics*, 2020, *1*, 1–23.

Cooley, Thomas F. and Gary D. Hansen, "The Inflation Tax in a Real Business Cycle Model," *American Economic Review*, 1989, *79* (4), 733–748.

Faig, Miquel and Belen Jerez, "A Theory of Commerce," *Journal of Economic Theory*, 2005, *122* (1), 60–99.

Geromichalos, Athanasios, Juan Manuel Licari, and José Suárez-Lledó, "Monetary Policy and Asset Prices," *Review of Economic Dynamics*, 2007, *10* (4), 761–779.

Khan, Aubhik and Julia K. Thomas, "Revisiting the Tale of Two Interest Rates With Endogenous Asset Market Segmentation," *Review of Economic Dynamics*, 2015, *18* (2), 243–268.

Kiyotaki, Nobuhiro and Randall Wright, "On Money as a Medium of Exchange," *Journal of Political Economy*, 1989, *97* (4), 927–954.

Kocherlakota, Narayana R., "Societal Benefits of Illiquid Bonds," *Journal of Economic Theory*, 2003, *108* (2), 179–193.

Lagos, Ricardo and Randall Wright, "A Unified Framework for Monetary Theory and Policy Analysis," *Journal of Political Economy*, 2005, *113* (3), 463–484.

_ , Guillaume Rocheteau, and Randall Wright, "Liquidity: A New Monetarist Perspective," *Journal of Economic Literature*, 2017, *55* (2), 371–440.

Lucas, Robert E., Jr., "Expectations and the Neutrality of Money," *Journal of Economic Theory*, 1972, *4* (2), 103–124.

_ , "Money in a Theory of Finance," *Carnegie-Rochester Conference Series on Public Policy*, 1984, *21*, 9–46.

_ , "Liquidity and Interest Rates," *Journal of Economic Theory*, 1990, *50* (2), 237–264.

_ , "Inflation and Welfare," *Econometrica*, 2000, *68* (2), 247–274.

Menzio, Guido, Shouyong Shi, and Hongfei Sun, "A Monetary Theory with Non-Degenerate Distributions," *Journal of Economic Theory*, 2013, *148* (6), 2266–2312.

Molico, Miguel, "The Distribution Of Money And Prices In Search Equilibrium," *International Economic Review*, 2006, *47* (3), 701–722.

Rocheteau, Guillaume and Antonio Rodriguez-Lopez, "Liquidity Provision, Interest Rates, and Unemployment," *Journal of Monetary Economics*, 2014, *65* (C), 80–101.

_ and Randall Wright, "Money in Search Equilibrium, in Competitive Equilibrium, and in Competitive Search Equilibrium," *Econometrica*, 2005, *73* (1), 175–202.

_ , Pierre-Olivier Weill, and Tsz-Nga Wong, "Working Through the Distribution: Money in the Short and Long Run," NBER Working Papers 21779, National Bureau of Economic Research, 2015.

_ , Randall Wright, and Sylvia Xiaolin Xiao, "Open Market Operations," *Journal of Monetary Economics*, 2018, *98*, 114–128.

Sargent, Thomas J. and Bruce D. Smith, "Irrelevance of Open Market Operations in Some Economies with Government Currency Being Dominated in Rate of Return," *American Economic Review*, 1987, *77* (1), 78–92.

— **and Neil Wallace**, "Some Unpleasant Monetarist Arithmetic," Quarterly Review Fall, Federal Reserve Bank of Minneapolis, 1981.

Shi, Shouyong, "A Divisible Search Model of Fiat Money," *Econometrica*, 1997, *65* (1), 75–102.

— , "Efficiency Improvement from Restricting the Liquidity of Nominal Bonds," *Journal of Monetary Economics*, 2008, *55* (6), 1025–1037.

Sterk, Vincent and Silvana Tenreyro, "The Transmission of Monetary Policy Through Redistributions and Durable Purchases," *Journal of Monetary Economics*, 2018, *99*, 124–137.

Trejos, Alberto and Randall Wright, "Search, Bargaining, Money, and Prices," *Journal of Political Economy*, 1995, *103* (1), 118–141.

Wallace, Neil, "A Modigliani-Miller Theorem for Open-Market Operations," *American Economic Review*, 1981, *71* (3), 267–274.

Williamson, Stephen D., "Liquidity, Monetary Policy, and the Financial Crisis: A New Monetarist Approach," *American Economic Review*, 2012, *102* (6), 2570–2605.

A Appendix

A.1 Solution Method

The numerical method employed here to solve for the stationary equilibrium has not, to my knowledge, been used in the literature of search-theoretic models of money. Although most papers in this area use assumptions that guarantee analytical tractability, the few that implement numerical methods to account for the intertemporal heterogeneity in money holdings employ solution strategies that are costly in terms of efficiency, that may sacrifice accuracy, and that are not easily extended to higher dimensions. For example, the numerical method in Molico (2006) and Chiu and Molico (2010, 2011) require simulating matches between agents and then performing a kernel density estimation of the distribution of money with these observations. Given that we must consider every possible meeting, the size of the simulation has to be large enough, so we have several observations for each match. This also implies that the required number of simulated matches increases with the dimension of the portfolio, as the number of possible matches may increase. In addition to this, having to use kernel estimation may imply losing accuracy in the estimated distribution.

My method solves iteratively for the agents' decision rules, the terms of trade in decentralized trading, the level of lump-sum transfers consistent with the government balance, and the prices that clear the markets of bonds and money. We start by solving the agents' value functions in (9) and (10). This requires updating $V(m, a; F, \theta)$ and $W(x; G, \theta)$ one after the other. Doing this for W in the decentralized market (DM) is standard, as long as we have how to compute the continuation value V. However, when computing the value function in the DM, we need to know the entire distribution $F(m, a) : \mathcal{M} \times \mathcal{A} \to \mathbb{R}$ and the terms of trade for each possible meeting between agents with (m_b, a_b) and (m_s, a_s), where $(m_i, a_i) \in \mathcal{M} \times \mathcal{A}$ for $i = b, s$. Therefore, we must start not only with some guess for W, as in any value function iteration approach, but also with guesses for $F(m, a)$ and $\langle Q, D \rangle = \langle q(m_b, a_b, m_s, a_s; F, \theta), d(m_b, a_b, m_s, a_s; F, \theta) \rangle$. To do this, we discretize our state space $\mathcal{M} \times \mathcal{A}$, so that $F(m, a)$ and $\langle Q, D \rangle$, as well as $V(m, a; F, \theta)$, are defined over a finite number of points. Similarly, when evaluating $W(x; G, \theta) : \mathcal{X} \to \mathbb{R}$, we need to discretize $\mathcal{X}$. Given this, we can solve the stationary equilibrium of this economy as follows:

1. Solve for V and W given some $F(m, a)$, some terms of trade $\langle Q, D \rangle$, a vector of prices (ϕ_m, ϕ_a), and some guess for the transfers, τ. As a byproduct of this step, we should have the decision rules for $m' = g_m(x; G, \theta)$ and $a' = g_a(x; G, \theta)$ in the

centralized market (CM).

There are several valid approaches to this. In particular, I use piece-wise cubic splines when evaluating continuation values that depend on W, solve the non-linear labor-leisure condition using the bisection method, and employ Howard's improvement algorithm to speed up the solution of the CM problem in (10)-(12).

2. Given $F(m, a)$ and the current values in $\langle Q, D \rangle$, compute $G(m, a)$. Likewise, using $G(m, a)$ and the decision rules for m' and a', compute $F'(m, a)$. Repeatedly update F and G until these distributions converge.

3. Verify if markets for money and bonds are clearing, i.e., check if Equations (5) and (6) hold. If not, adjust the prices (ϕ_m, ϕ_a) and return to Step 1. Repeat this process until clearing both markets. Of course, when returning to Step 1, use the most recently computed value for F.

   When adjusting prices, an alternative is to use the excess demand functions of bonds and money. I do so using a partial updating approach.

4. Check if the government balance in Equation (2) is satisfied. In case it is not, adjust the tax rate in Equation (3) using the most recent value for ϕ_a and go back to Step 1.

5. Finally, update the terms of trade $\langle Q, D \rangle$. The key input here is the value function W that is required to compute the continuation values in the problem in (7)-(8). As in Step 1, I use piece-wise cubic splines on W. Check if the newly calculated terms of trade are close enough to their previous values. If not, return to Step 1. If so, we are done.

A.2 Additional Figures and Tables

Table A1: Elasticity of Model Moments with Respect to Parameters

Parameter	α	B	ν	χ	κ
Ratio of bonds to money	0.70	0.00	-0.93	1.06	-1.61
Total output	0.41	0.00	-0.57	0.33	-0.72
Output in DM	1.04	0.00	-1.43	0.39	-0.83
Output in CM	0.00	0.00	-0.01	0.30	-0.66
Std. dev. of money in DM	1.28	0.00	0.93	2.01	1.54
Std. dev. of money in CM	0.60	0.00	0.08	0.42	0.08
Std. dev. of bonds	0.86	0.00	-0.94	1.09	-1.65
Average DM price	-0.06	0.35	1.75	0.32	0.05
Std. dev. of DM price	0.54	0.35	0.88	1.32	-0.52
Velocity	0.40	0.00	0.93	0.31	0.15

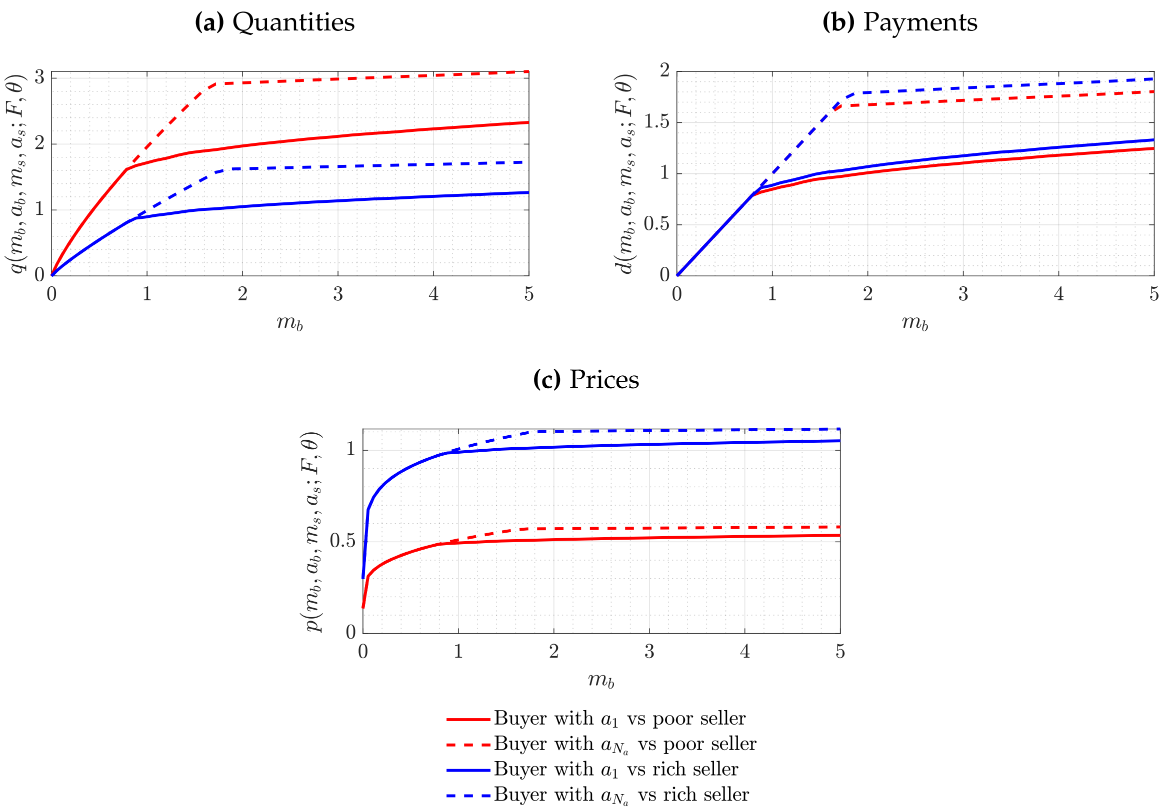
Notes: Elasticities of model moments (rows) with respect to parameters (columns) around the calibrated parameters for the baseline stationary equilibrium. α is the probability of being matched in the DM, B and ν are the scale and curvature parameters of the cost of working in the DM, χ is the inverse of the Frisch elasticity of labor supply in the CM, and κ is the scale parameter for the disutility of labor in the CM.

Table A2: Long-Run Distributional Effects of an Open Market Operation from $i = 4.9\%$ to $i = 3.9\%$

Percentile	≤ 10	**10-25**	**25-50**	**50-75**	**75-90**	≥ 90
Consumption (CM)	0.41	-2.13	-0.26	0.29	1.01	1.43
Labor	-0.22	1.15	0.15	-0.15	-0.50	-0.70
Money holdings	-4.75	-3.35	-0.22	0.26	1.90	2.97
Bond holdings	–	-82.65	-29.95	-20.79	-18.12	-16.56
Consumption (DM)	-3.29	-3.36	-0.96	-0.81	0.31	1.32

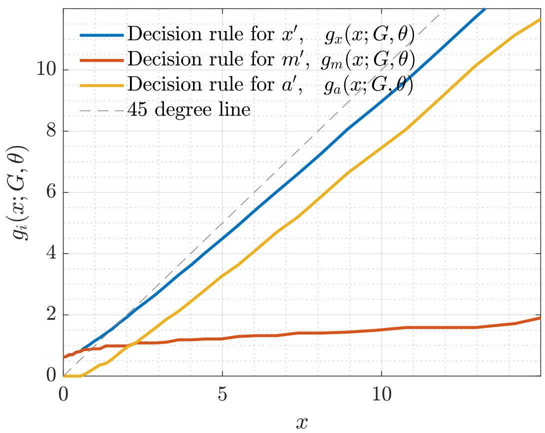
Notes: Percent changes by percentile groups across stationary equilibria with different targeted nominal interest rates. Changes are with respect to the baseline economy with a 4.9% nominal interest rate. Agents in the lowest decile have zero bond holdings in both stationary equilibria.

Figure A1: More on Terms of Trade: Quantities, Monetary Payments, and Prices as Function of Money from Buyer, m_b



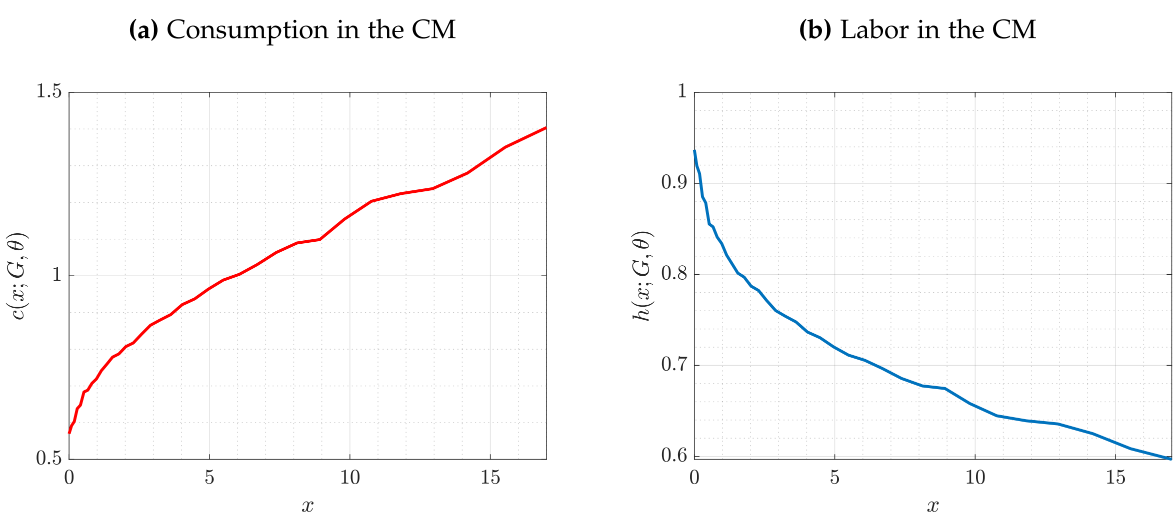
Notes: Terms of trade as a function of buyer's money holdings. This includes the outcomes for combinations of meetings between either bond-poor ($a_b = a_1$) or bond-rich ($a_b = a_N$) buyers, and either asset-poor ($z_s = (m_1, a_1)$) or asset-rich ($z_s = (m_{N_m}, a_{N_a})$) sellers. In the numerical solution, m_1 and a_1 are the lowest possible value for money and bond holdings, while m_{N_m} and a_{N_a} are the highest. As a reference, the efficient level of trading for the baseline calibration is $q^* = 1.98$.

Figure A2: Decision Rules for Assets



Notes: Decision rules for money, $m' = g_m(x; G, \theta)$; bonds, $a' = g_a(x; G, \theta)$; and the implied total wealth, $x' = m' + a' = g_x(x; G; \theta)$, as a function of wealth, x, in the centralized market.

Figure A3: Decision Rules for Consumption and Labor in the CM



Notes: Decision rules for consumption, $c = c(x; G, \theta)$, and labor, $h' = h(x; G, \theta)$, as a function of wealth, x, in the centralized market.

B Persistent, Nondegenerate Distribution with Only Money

Below I present a version of the model where money is the only asset and, hence, monetary policy does not operate through OMOs but via lump-sum transfers to agents. This model economy resembles the one in Lagos and Wright (2005), but with non-quasi-linear preferences. This simplified model illustrates how much heterogeneity this environment can generate as it moves away from having a quasi-linear utility function in the CM and we add more curvature to the labor supply. It is also helpful to explain the numerical solution method in a simpler manner.

As in the full model, each period is divided into decentralized and centralized markets operating sequentially. Matching in the DM occurs as in the full model, with agents trading using only money and terms of trade being determined by a "take it or leave it" offer of the buyer to the seller. The agents' period utility function is as in the full model, with U not restricted to be linear in h. The production technologies in both the decentralized and centralized markets operate as in the full model.

The monetary authority controls the supply of the only available asset to agents: fiat money, M. The supply of money grows at the rate μ. At the beginning of every CM, the monetary authority transfers μM to every agent in the economy before the CM begins. This implies that the stock of aggregate nominal money evolves according to:

$$M' = (1+\mu)\, M, \tag{B.1}$$

where M' is the nominal money supply at the middle of the current period, and therefore, at the beginning of next period. Individual nominal money holdings are normalized with respect to the beginning of the period money supply, M. This means that if an agent holds $\hat{m}$ units of nominal money, then they have $m = \hat{m}/M$ units of relative money holdings. Note that money injections μM are equivalent to μ units of relative money.

Following a similar notation as before, the distribution of agents over m are summarized by the probability measures $F(m)$ and $G(m)$ for the DM and CM, respectively. These distributions evolve according to $G = \Gamma_G(F, \mu)$ and $F' = \Gamma_F(G, \mu)$. Given this, and the policy implemented by the monetary authority, in equilibrium we have:

$$\int_{\mathcal{M}} m dF(m) = 1 \tag{B.2}$$

$$\int_{\mathcal{M}} m dG(m) = 1 + \mu. \tag{B.3}$$

Let $V(m; F, \mu)$ and $W(m; G, \mu)$ be the value functions at the beginning of the DM and CM, respectively. As noted before, without loss of generality, we assume that there are not double-coincidence meetings. In single-coincidence meetings, a buyer and a seller with relative money holdings m_b and m_s are matched. The buyer makes a "take it or leave it" offer to the seller: the buyer offers to buy q of the differentiated good at the price d. The terms of such an offer are determined according to the following problem:

$$\max_{q,d} \; u(q) + W(m_b + \mu - d; G, \mu) \tag{B.4}$$

subject to the seller's participation constraint

$$-v(q) + W(m_s + \mu + d; G, \mu) \geq W(m_s + \mu; G, \mu) \tag{B.5}$$

to the law of motion of the distribution, $G(m) = \Gamma_G(F(m), \mu)$, and to $0 \leq d \leq m_b$, $q \geq 0$. Note that the continuation values of buyers and sellers take into account their money holdings when exiting the current DM and the transfer they will receive at the beginning of the next CM. For each meeting (m_b, m_s), and given an aggregate state $\{F(m), \mu\}$, we have that the terms of trade $q(m_b, m_s; F, \mu)$ and $d(m_b, m_s; F, \mu)$ solve the problem stated above.

In this context, the expected lifetime utility at the beginning of the DM of an agent with money holdings m, i.e., before knowing if they are matched or not, and before knowing their role in an eventual match, is given by the following functional equation:

$$\begin{aligned} V(m; F, \mu) = & \frac{\alpha}{2} \int_{\mathcal{M}} \{ \;\; u(q(m, m_s; F, \mu)) + W(m + \mu - d(m, m_s; F, \mu); G, \mu)\} \, dF(m_s) \\ & + \frac{\alpha}{2} \int_{\mathcal{M}} \{-v(q(m_b, m; F, \mu)) + W(m + \mu + d(m_b, m; F, \mu); G, \mu)\} \, dF(m_b) \\ & + (1 - \alpha) W(m + \mu; G, \mu) . \end{aligned} \tag{B.6}$$

Let $\tilde{m}$ denote relative money holdings at the beginning of the CM. Money holdings at the beginning of the CM are those at the end of the previous DM, plus lump-sum transfers of money. For an agent with relative money holdings $\tilde{m}$, the lifetime value of their utility at the beginning of the CM is given by:

$$W(\tilde{m}; G, \mu) = \max_{c,h,m'} \left\{ U(c, h) + \beta V\left(m'; F', \mu\right) \right\} \tag{B.7}$$

subject to the budget constraint

$$c = h + \phi(G, \mu)\left[\tilde{m} - (1+\mu)\, m'\right] \tag{B.8}$$

and to the mapping of $G(m)$ into $F'(m)$

$$F'(m) = \Gamma_F(G(m), \mu). \tag{B.9}$$

For the utility and cost functions in the DM, as in the full model, I follow Lagos and Wright (2005). However, for the CM I drop the quasi-linearity assumption. In particular, the utility function in the CM is concave in both consumption and leisure and given by:

$$U(c, h) = \log c - \kappa \frac{h^{1+\chi}}{1+\chi}, \tag{B.10}$$

where χ is the inverse of the Frisch elasticity of labor supply and $\kappa > 0$ a scale parameter. Note that (B.10) nests the quasi-linear utility function in Lagos and Wright (2005) when $\chi = 0$ and $\kappa = 1$.

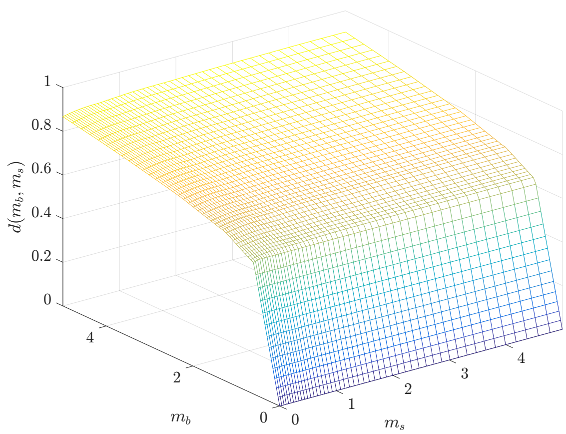
The model is parametrized as closely as possible to Lagos and Wright (2005) for comparison purposes. One model period is equivalent to one year. The discount factor is set to generate a real interest rate of 4 percent, and the money growth rate is consistent with an inflation rate, in the stationary equilibrium, of 2 percent. The probability of being matched in the DM, α, is set to 1 to minimize the role of the search frictions. The inverse of the Frisch elasticity is chosen to be 0.5. I select B and κ (scale parameters), as well as ν (the curvature parameter that pins down the cost of producing in the DM) to reproduce the following three targets: (1) the size of the CM to be 10 percent of total output, (2) a velocity of money of 2, and (3) an average markup in DM meetings of 30 percent (Faig and Jerez, 2005). Finally, the parameter that governs the curvature of the utility function in the DM, η, is set close to 1, so that the utility function is close to being logarithmic.

B.1 Stationary Equilibrium

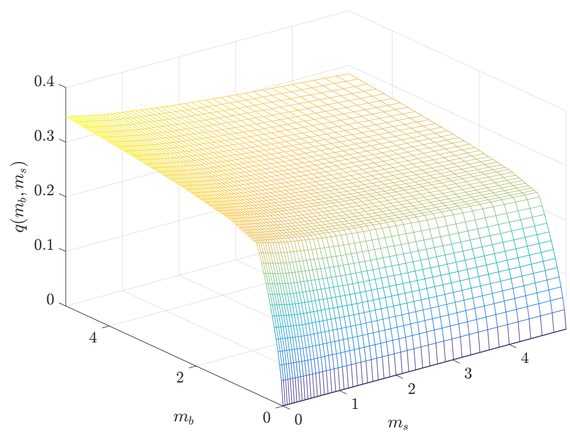
In a stationary equilibrium, the distribution over relative money holdings remains constant, i.e., $F = \Gamma_F(\Gamma_G(F, \mu), \mu)$. This, in turn, implies $\phi = \phi'$ and requires the growth rate of prices, π, to be equal to the growth rate of money, μ. In contrast to models like Shi (1997) or Lagos and Wright (2005), the distribution of money holdings is not necessarily degenerate across periods.

Figure B4: Terms of Trade in the Decentralized Market

(a) Payments



(b) Quantities



(c) Price per unit

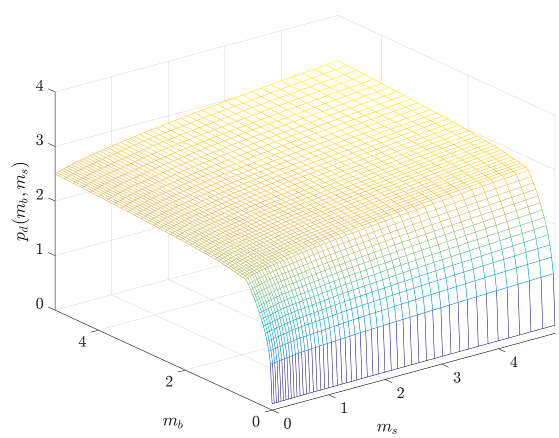
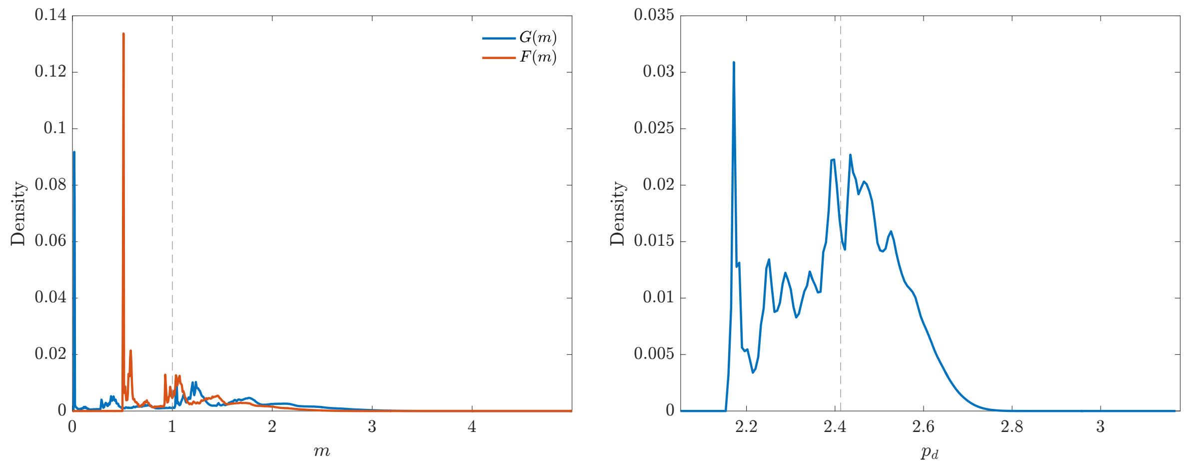


Figure B4 shows the stationary terms of trade $d(m_b, m_s; F, \mu)$ and $q(m_b, m_s; F, \mu)$. As opposed to models where preferences in the CM are linear in labor, here, as in the full model, terms of trade depend not only on the buyer's money holdings but also on the seller's. For a fixed level of m_s, larger values of m_b imply that the buyer can obtain more of the differentiated good but at a higher total cost. On the other hand, keeping fixed m_b, a seller with larger money balances requires a higher payment for lower quantities of the good he produces. This is because of the seller's continuation value (captured by the function W), which is now concave in m. Consequently, as Figure B4(c) shows, the price per unit of good exchanged, $p_d(m_b, m_s) = d(m_b, m_s) / q(m_b, m_s)$, mainly responds to the seller's wealth.

Figure B5: Distributions at the Stationary Equilibrium

(a) Distribution of Money Holdings **(b)** Distribution of Prices



The chance of having different types of meetings with heterogeneity in both q and d, plus the fact that preferences in the CM are not quasi-linear, generates a nondegenerate distribution of money holding, as shown in Figure B5(a). In that figure, we have the distributions $F(m)$ and $G(m)$. The distribution entering the CM, unsurprisingly, is more dispersed than the distribution at the beginning of the DM, reflecting the heterogeneity-inducing effect of decentralized trading. We can also observe that, after trading in the DM, some agents end up with very little money holdings (the money injections prevent them from reaching $m = 0$). This is the result of a combination of sequences of meetings with sellers that make the buyer deplete her money holdings. However, even for this "unlucky" type of agent, the CM acts as an insurance market that lets her replenish, at least partially, her liquidity for the next round of decentralized trading. Note that some agents reaching the lower end of the distribution may experience similar matches in the future as they move away from $m = 0$. This explains the spikes in $F(m)$ and why they get diluted for higher values of money holdings.

Figure B5(b), shows the observed distribution of prices. This distribution comes from taking into account the likelihood of all possible meetings implied by the distribution $F(m)$ and the prices $p_d(m_b, m_s)$ at each one of them. In this context, the model is able to produce a nondegenerate distribution of prices that reflects the differences in assets between different agents in this economy.

B.2 The Role of the Elasticity of Labor Supply

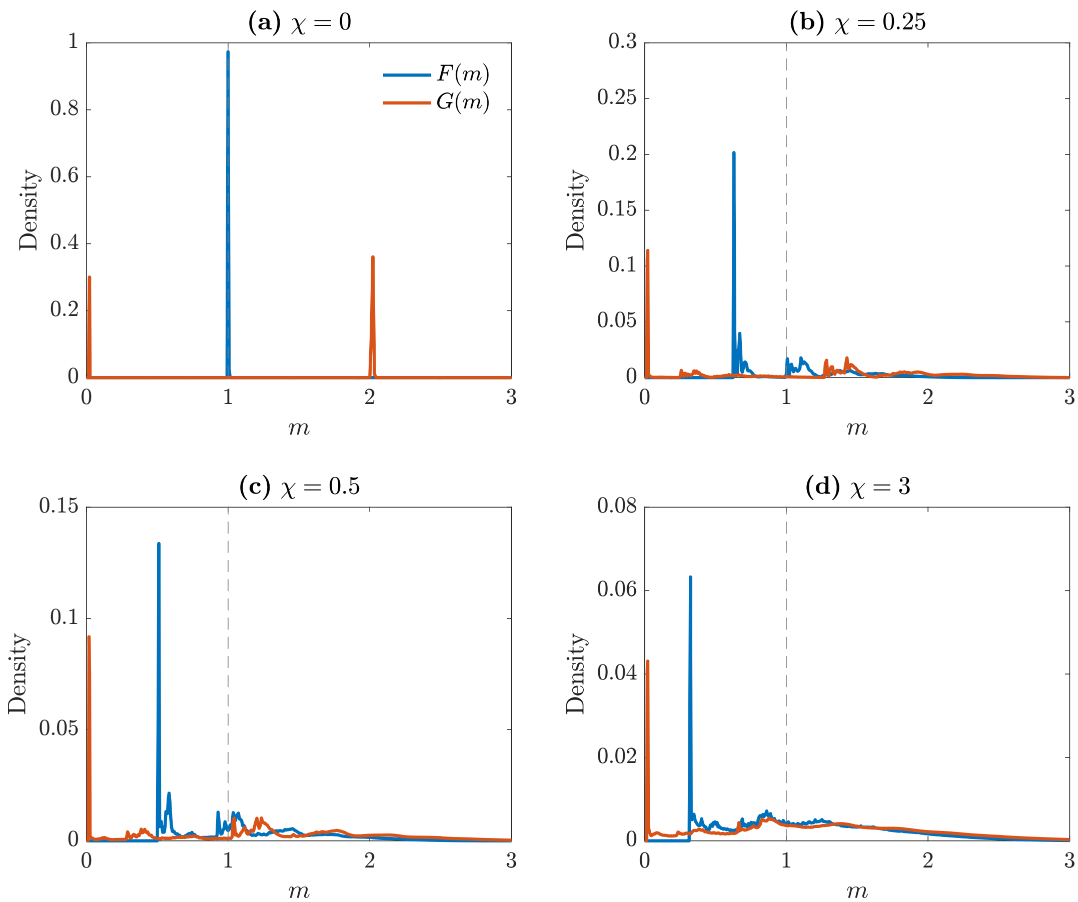
The elasticity of labor supply is crucial in determining the extent to which agents can reshuffle their money holdings after trading in the DM. The higher this elasticity, the easier for agents to offset their previous idiosyncratic trading histories. In Lagos and Wright (2005), labor supply is infinitely elastic, so agents are able to completely undo in the CM whatever happened in their previous DM meeting. In this model, having a perfectly elastic labor supply is equivalent to $\chi = 0$.

Figure B6 shows the stationary distributions of money at the beginning of the DM and CM for different values of χ, the inverse of the Frisch elasticity of labor. When $\chi = 0$, we obtain a degenerate distribution of money in the DM (blue line in panel (a) of Figure B6): all agents are holding the same amount of relative units of money. As these agents trade in decentralized meetings, buyers deplete their money holdings, while sellers end up with about double what they initially had (red line).[14] However, the fact that, for this case, labor is perfectly elastic allows agents to reshuffle back their money holdings completely. The decision rules for labor and money holdings in Figure B7(a) and Figure B7(b) show that labor supply changes linearly with the individual state, m, so that agents can exactly exit the period with 1 unit of relative money. This is a consequence of the absence of wealth effects under quasi-linear preferences.

When the Frisch elasticity of labor supply ($1/\chi$) becomes finite, the inability of agents to offset the effects of their previous trading history results in a nondegenerate distribution of money in the DM. In fact, as the elasticity of labor supply decreases, we obtain distributions of money holdings with less concentration.

[14] When $\chi = 0$, the CM distribution is split equally between two points, each one with mass 0.5. In Figure B6, this does not appear to happen because of a small numerical error. These two points are not necessarily in the grid $\mathcal{M}$, so that the mass of agents gets distributed between their adjacent points. In any case, adding the masses at those adjacent points gives exactly 0.5 for each one.

Figure B6: Distributions of Money Balances at the Stationary Equilibrium for Different Levels of the Frisch Elasticity of Labor



Notes: Stationary distributions of money holdings at the beginning of the decentralized market, $F(m)$, and the beginning of the centralized market, $G(m)$, for different values of the Frisch elasticity of labor supply, $1/\chi$.

Figure B7: Decision Rules in the Centralized Market at the Stationary Equilibrium for Different Levels of the Frisch Elasticity of Labor

(a) Decision Rules for Labor **(b)** Decision Rules for Money Holdings