Risk Premia and the Real Effects of Money

Sebastian Di Tella
Stanford University

May 2019

Abstract

This paper proposes a flexible-price theory of the role of money in an economy with incomplete idiosyncratic risk sharing. When the risk premium goes up, money provides a safe store of value that prevents interest rates from falling, depressing investment. Investment is too high during booms when risk is low, and too low during slumps when risk is high. Monetary policy is ineffective—money is superneutral and Ricardian equivalence holds. The optimal allocation requires the Friedman rule and a tax/subsidy on capital. The real effects of money survive even in the cashless limit.

1 Introduction

This paper studies the role of money in an economy with incomplete idiosyncratic risk sharing. I show that money can play a central role in how the economy reacts to an increase in the risk premium, even if prices are completely flexible. During downturns idiosyncratic risk goes up, raising the risk premium and making risky capital less attractive. Without money, real interest rates would fall and keep investment stable. Money prevents equilibrium interest rates from falling, depressing investment.

The baseline model is a simple AK growth model with log utility over consumption and money, and incomplete idiosyncratic risk sharing. While the presence of money has important real effects, money is superneutral and Ricardian equivalence holds. The competitive equilibrium is not efficient, however. Investment is too high during booms when risk is low, but too low during slumps when risk is high. In contrast, without money investment is always too high. Implementing the optimal allocation requires the Friedman rule and a tax or subsidy on capital.

To understand the role of money, it's useful to proceed in two steps. First, for a given level of risk, money provides a safe store of value that improves idiosyncratic risk sharing and weakens agents' precautionary saving motive relative to the risk premium on capital. This keeps real interest rates high and investment depressed relative to the non-monetary economy. Second, the value of money increases endogenously with risk. The value of money is the present value of expenditures on liquidity services, and it becomes very large when real interest rates fall. In particular, if risk is high enough the real interest rate can be very negative without money, but must remain above the growth rate if there is money. The value of money endogenously grows, raising the equilibrium real interest rate and depressing investment until this condition is satisfied.

Money provides a safe store of value because its liquidity premium makes it effectively in positive net supply. I show that safe private and public debt performs the same role as money only to the extent that it has a liquidity premium (such as deposits or short-term Treasuries).[1] To see why, notice that safe assets without a liquidity premium (private or public) must be backed by payments with equal present value. Agents own the assets but also the liabilities, so the net value is zero. But the value of liquid assets, net of the value of the payments backing them, is equal to the present value of their liquidity premium. This is what allows them to serve as a store of value and improve risk sharing in general equilibrium.

There are many real assets that have positive net value, such as capital, housing, or land. But the starting point of this paper is that real investments are risky and idiosyncratic risk sharing is incomplete. For example, the value of a particular house or plot of land has significant idiosyncratic risk that can't be fully shared. There are also safe financial assets, such as AAA corporate debt, but their net value is zero (someone owes the asset). Safe assets with a liquidity premium have the rare combination of safety and positive net value that allow them to function as a safe store of value.

I also allow agents to issue outside equity that can be diversified into a safe equity index. While issuing outside equity improves risk sharing, it does not perform the same role as money and other safe liquid assets. Ultimately, equity is also in zero net supply (someone issues the equity). Specifically, with outside equity but without money, after an increase in idiosyncratic risk real interest rates fall and investment remains stable.

How quantitatively important is the role of liquid assets as a safe store of value? The net value of liquid assets is equal to the present value of expenditures on liquidity services. During periods when real interest rates are high this value is relatively small, close to the expenditure share on liquidity services (around 1.7%) of total wealth. But when real interest rates become persistently very low, such as in the aftermath of the 2008 financial crisis, the net value of liquid assets can become very large. In fact, the real effects of money survive even in the cashless limit where expenditures on liquidity services vanish, and are robust to different specifications of money demand.

The inefficiency in this economy comes from the presence of hidden trade.[2] I use a mechanism-design approach to study optimal policy. I microfound the incomplete idiosyncratic risk sharing with a fund diversion problem with hidden trade. The competitive equilibrium is the outcome of allowing agents to write privately optimal long-term contracts in a competitive market. I then characterize the optimal allocation by a planner who faces the same environment, and show how it can be implemented with a tax or subsidy on capital.

We can understand the inefficiency in terms of agents' precautionary saving motive and the risk premium of capital. On the one hand, the precautionary motive leads to overinvestment as agents attempt to save to self insure. On the other hand, an excessively large risk premium leads

[1] See Krishnamurthy and Vissing-Jorgensen (2012).

[2] See Farhi et al. (2009), Kehoe and Levine (1993), Di Tella (2016).

to underinvestment. Money improves risk sharing and weakens the precautionary motive relative to the risk premium. As a result, when the value of money is small (when idiosyncratic risk is low) the precautionary motive dominates, and the competitive equilibrium features too much investment and too little risk sharing. But when the value of money is large (when idiosyncratic risk is high), the excessive risk premium dominates and the competitive equilibrium features too little investment and too much risk sharing.

The model is driven by countercyclical idiosyncratic risk shocks for the sake of concreteness. But an increase in risk aversion is mathematically equivalent; it will also raise the risk premium and precautionary motives. Higher risk aversion can represent wealth redistribution from risk tolerant to risk averse agents after bad shocks (see Longstaff and Wang (2012) and Gârleanu and Panageas (2015)) or weak balance sheets of specialized agents who carry out risky investments (see He and Krishnamurthy (2013) and He et al. (2015)). It can also capture habits (Campbell and Cochrane (1999)) or higher ambiguity aversion after shocks that upend agents' understanding of the economy (see Barillas et al. (2009)). Here I focus on simple countercyclical risk shocks with homogenous agents, but these are potential avenues for future research.

I first study a simple stationary model and consider unanticipated and permanent shocks to idiosyncratic risk (equivalent to comparative statics across balanced growth paths) in section 2. I also show that the real effects of money survive in the cashless limit, and are robust to alternative specifications of money demand. In section 3 I characterize the optimal allocation. The stationary environment captures most of the economic intuition and can be solved with pencil and paper. I then introduce aggregate risk shocks in a dynamic model in section 4 and characterize the competitive equilibrium as the solution to a simple ODE. Section 5 discusses the link to a bubble theory of money, the relationship between the mechanism in this paper and sticky-price models of the zero lower bound, and several extensions. In the Appendix, I also solve the model with Epstein-Zin preferences. The Online Appendix has the technical details of the contractual environment.

Literature review. The mainstream view of the role of money focuses on the effect of nominal rigidities in the context of New Keynesian models. If there is money in the economy the nominal interest rate cannot be negative. So if the natural interest rate (the real interest rate with flexible prices) is very negative, the central bank must either abandon its inflation target or allow the economy to operate with an output gap (or both).[3] In contrast, in this paper prices are flexible, and the zero lower bound is not binding and does not play any role. Depressed investment does not reflect an output gap, but rather the equilibrium real effects of money.

The results in this paper have an important takeaway for New Keynesian models of the zero lower bound. Introducing money into an economy doesn't just place a lower bound on interest rates—it also raises the natural interest rate. It's possible for the natural interest rate to be negative without money, but positive once money is introduced, so that the zero lower bound is not binding.

[3] Krugman et al. (1998), Eggertsson et al. (2003), Werning (2011), Eggertsson and Woodford (2004), Eggertsson and Krugman (2012), Svensson (2000), Caballero and Simsek (2017).

Buera and Nicolini (2014) provide a flexible-price model where money has real effects at the zero lower bound, based on borrowing constraints and lack of Ricardian equivalence. Aiyagari and McGrattan (1998) study the role of government debt in a model with uninsurable labor income and binding borrowing constraints. In contrast, here Ricardian equivalence holds (agents have the natural borrowing limit), the zero lower bound is not binding, and money is superneutral. Changing the amount of government debt can only affect the liquidity premium on government debt and other assets, but not the real side of the economy. It is easy to break Ricardian equivalence and superneutrality, but they are useful theoretical benchmarks that highlight that the mechanism does not hinge on a fiscal channel.

The liquidity premium is the focus of a large literature that microfounds the role of money as a means of exchange in a search-theoretic framework.[4] Here I use money in the utility function as a simple and transparent way to introduce money into the economy. I also show the main results are robust to a cash-in-advance specification. The purpose of this paper is not to provide a new explanation for why people hold money in equilibrium, but rather to understand how money can have real effects on the equilibrium. However, a more microfounded account of the liquidity premium can help understand how it is affected by aggregate shocks and policy interventions. In this same line, a classic question in monetary economics concerns the role of inflation on investment and growth.[5] In contrast, here money is superneutral, so inflation doesn't have real effects.

There is also a large literature modeling money as a bubble in the context of OLG or incomplete risk sharing models.[6] The closest paper is Brunnermeier and Sannikov (2016b), who use a similar environment with incomplete idiosyncratic risk sharing to study the optimal inflation rate.[7] An important contribution of that paper is to develop a version of the Bewley (1980) model of bubble money that is tractable and yields closed-form solutions. They find that countries with high risk should have a higher inflation rate. In contrast, here bubbles are ruled out, money is superneutral, and the focus is on how money can have real effects.[8] I study some of the differences and similarities of the bubble and liquidity views of money in section 5.

Many papers highlight the role of risk or uncertainty shocks both in macro and finance.[9] The setting here is closest to Di Tella (2017), who shows that risk shocks that increase idiosyncratic risk can help explain the concentration of aggregate risk on the balance sheets of financial intermediaries and create financial crises. Here I remove intermediaries and introduce money.

Cochrane (2011) highlights the role of time-varying risk premia in asset prices, and therefore investment. The driving force here is a time-varying idiosyncratic risk premium. But a high risk premium is not enough to depress investment. If the real interest rate drops enough it can stabilize

[4] See Kiyotaki and Wright (1993), Lagos and Wright (2005), Aiyagari and Wallace (1991), Shi (1997).

[5] See Stockman (1981), Sidrauski (1967), Aruoba et al. (2011), Dotsey and Sarte (2000), Tobin (1965).

[6] See Samuelson (1958), Bewley (1980), Diamond (1965), Tirole (1985), Asriyan et al. (2016), Santos and Woodford (1997)

[7] Brunnermeier and Sannikov (2016a) use a similar environment but focus on the role of financial intermediaries.

[8] In their model money is a bubble and is introduced proportionally to wealth, so higher inflation acts as a subsidy to saving. Here bubbles are explicitly ruled out and money is introduced in a lump-sum, non-distortionary way.

[9] Bloom (2009), Bloom et al. (2012), Campbell et al. (2001), Bansal and Yaron (2004), Bansal et al. (2014), Campbell et al. (2012), Christiano et al. (2014).

asset prices and investment. This paper provides a theory of why the equilibrium real interest rate will not drop enough, so that a high risk premium will be reflected in lower investment.

The contractual environment microfounding the incomplete idiosyncratic risk sharing with a fund diversion problem with hidden trade is based on Di Tella and Sannikov (2016), who study a more general environment.[10] Di Tella (2016) uses a similar contractual environment to study optimal financial regulation, but does not allow hidden savings or investment. Instead, it focuses on the externality produced by hidden trade in capital assets by financial intermediaries. That externality is absent in this paper because the price of capital is always one (capital and consumption goods can be transformed one-to-one).

2 Baseline model

In this section I introduce the baseline stationary model. It's a simple AK growth model with money in the utility function and incomplete idiosyncratic risk sharing. The equilibrium is always a balanced growth path, and to keep things simple I will consider completely unexpected and permanent risk shocks that increase idiosyncratic risk (equivalent to comparative statics across balanced growth paths). In section 4 I will introduce the fully dynamic model with aggregate risk shocks.

2.1 Setting

The economy is populated by a continuum of agents with log preferences over consumption c and real money balances $m \equiv M/p$

$$U(c,m) = \mathbb{E}\left[\int_0^\infty e^{-\rho t}\big((1-\beta)\log c_t + \beta \log m_t\big)dt\right]$$

Money in the utility function is a simple and transparent way of introducing money in the economy. In section 2.6 I also solve the model with a cash-in-advance constraint and a more general CES utility function. As we'll see, what matters is that money has a liquidity premium.

Agents can continuously trade capital and use it to produce consumption $y_t = ak_t$, but it is exposed to idiosyncratic "quality of capital" shocks. The change in an agent's capital over a small period of time is

$$d\Delta^k_{i,t} = k_{i,t}\sigma dW_{i,t}$$

where $k_{i,t}$ is the agent's capital (a choice variable) and $W_{i,t}$ an idiosyncratic Brownian motion. Idiosyncratic risk σ is a constant here, but we will look at the effects of an unexpected and permanent change in σ (equivalent to comparative statics of the equilibrium with respect to changes in σ). This is meant to capture a shock that makes capital less attractive and drives up its risk premium. Later we will introduce a stochastic process for σ in a fully dynamic model.

[10]Cole and Kocherlakota (2001) study an environment with hidden savings and risky exogenous income, and find that the optimal contract is risk-free debt. Here we also have risky investment.

Idiosyncratic risk washes away in the aggregate, so the aggregate capital stock k_t evolves

$$dk_t = (x_t - \delta k_t)dt \tag{1}$$

where x_t is investment. The aggregate resource constraint is

$$c_t + x_t = ak_t \tag{2}$$

where c_t is aggregate consumption.

Money is printed by the government and transferred lump-sum to agents. In order to eliminate any fiscal policy, there are no taxes, government expenditures, or government debt; later I will introduce safe government debt and taxes. For now money is only currency, but later I will add deposits and liquid government bonds. The total money stock M_t evolves

$$\frac{dM_t}{M_t} = \mu_M dt$$

The central bank chooses μ_M endogenously to deliver a target inflation rate π. This means that in a balanced growth path $\mu_M = \pi + \text{growth rate}$.

Markets are incomplete in the sense that idiosyncratic risk cannot be shared. They are otherwise complete. Agents can continuously trade capital at equilibrium price $q_t = 1$ (consumption goods can be transformed one-to-one into capital goods, and the other way around) and debt with real interest rate $r_t = i_t - \pi$, where i_t is the nominal interest rate. There are no aggregate shocks for now; I will add them later and assume that markets are complete for aggregate shocks.

Total wealth is $w_t = k_t + m_t + h_t$, which includes the capitalized real value of future money transfers

$$h_t = \int_t^\infty e^{-\int_t^s r_u du} \frac{dM_s}{p_s} \tag{3}$$

The dynamic budget constraint for an agent is

$$dw_t = (r_t w_t + k_t \alpha_t - c_t - m_t i_t)dt + k_t \sigma dW_t \tag{4}$$

with solvency constraint $w_t \geq 0$, where $\alpha_t \equiv a - \delta - r_t$ is the excess return on capital. Each agent chooses a plan (c, m, k) to maximize utility $U(c, m)$ subject to the budget constraint (4).

Remark 1. As in Brunnermeier and Sannikov (2016b) and Angeletos (2006), this setting has several features that make it very tractable and easy to solve in closed form with pencil and paper. Uninsurable idiosyncratic risk comes from tradable capital, rather than non-tradable labor income. Together with homothetic preferences, this produces policy functions linear in wealth, which eliminate the need to keep track of the whole wealth distribution and yields closed-form expressions.[11]

[11] Angeletos (2006) does not have money. Brunnermeier and Sannikov (2016b) develop a tractable version of the Bewley (1980) model of bubble money. They introduce money proportionally to wealth. Here bubbles are explicitly ruled out and money is introduced lump-sum. I discuss the similarities and differences between the bubble and liquidity views of money in section 5.

Remark 2. Including future money transfers in the definition of wealth is equivalent to the natural borrowing limit.[12] The only friction in this economy is incomplete idiosyncratic risk sharing. As we'll see, money is superneutral, so there is no loss from setting the growth rate of money $\mu_M = 0$, so that $h_t = 0$ always. Allowing $\mu_M \neq 0$ and $h_t \neq 0$ is important to establish superneutrality and Ricardian equivalence.

2.2 Balanced growth path equilibrium

The competitive equilibrium is a balanced growth path (BGP). It is scale invariant to aggregate capital k_t, so we can normalize all variables by k_t; e.g. $\hat{m}_t = m_t/k_t$. A *Balanced Growth Path Equilibrium* consists of a real interest rate r, investment $\hat{x}$, and real money $\hat{m}$ satisfying

$$r = \rho + (\hat{x} - \delta) - \sigma_c^2 \qquad \text{Euler equation} \qquad (5)$$

$$r = a - \delta - \sigma_c \sigma \qquad \text{asset pricing} \qquad (6)$$

$$\sigma_c \equiv \frac{k_t}{k_t + m_t + h_t}\sigma = (1 - \lambda)\sigma \qquad \text{risk sharing} \qquad (7)$$

$$\lambda \equiv \frac{m_t + h_t}{k_t + m_t + h_t} = \frac{\rho\beta}{\rho - ((1 - \lambda)\sigma)^2} \qquad \text{value of liquidity} \qquad (8)$$

$$\hat{m} = \frac{\beta}{1 - \beta}\frac{a - \hat{x}}{r + \pi} \qquad \text{money} \qquad (9)$$

As well as $i = r + \pi > 0$ and $r > (\hat{x} - \delta)$. These last conditions make sure money demand is well defined and rule out bubbles.

Equation (5) is the usual Euler equation. $\hat{x} - \delta$ is the growth rate of the economy and therefore consumption, and σ_c^2 is the precautionary saving motive. The more risky consumption is, the more agents prefer to postpone consumption and save. Equation (6) is an asset pricing equation for capital. Agents can choose to invest their savings in a risk-free bond (in zero net supply) and earn r, or in capital and earn the marginal product net of depreciation $a - \delta$. The last term $\alpha = \sigma_c \sigma$ is the risk premium on capital. Because the idiosyncratic risk in capital cannot be shared, agents will only invest in capital if it yields a premium to compensate them.

Equation (7) is agents' exposure to idiosyncratic risk. Because of homothetic preferences each agent consumes proportionally to his wealth, and his exposure to idiosyncratic risk comes from his investment in capital. In equilibrium, the portfolio weight on capital is $k_t/w_t = k_t/(k_t + m_t + h_t) = (1 - \lambda)$ where we define $\lambda \equiv (m_t + h_t)/w_t$ as the share of wealth in money (present and future). λ captures the value of liquidity in the economy, and (8) gives us an equation for λ in terms of parameters. Finally, (9) is an expression for real money balances. Because of the log preferences agents devote a fraction β of expenditures to liquidity and $1 - \beta$ to consumption. Using $i = r + \pi$ and the resource constraint (2), we obtain (9).

[12] Define current assets as $\tilde{w}_t = k_t + m_t + d_t$, where d_t is risk-free debt (in zero net supply). The dynamic budget constraint is $d\tilde{w}_t = (d_t r_t + k_t(a - \delta) + \tau_{Mt} - m_t\pi - c_t)dt + k_t\sigma dW_t$, where τ_{Mt} are the money transfers, and the natural debt limit $\underline{\tilde{w}}_t = -h_t$, so that $\tilde{w}_t \geq \underline{\tilde{w}}_t$. This is equivalent to (4) with $w_t = \tilde{w}_t + h_t \geq 0$.

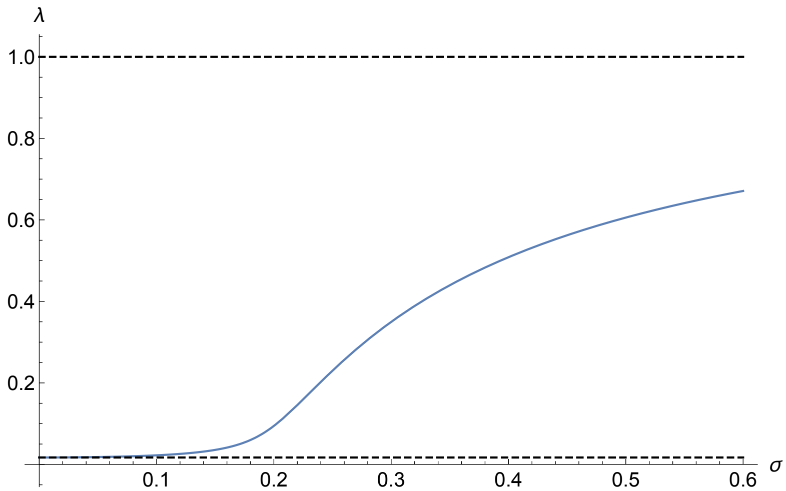

Figure 1: The value of liquidity λ as a function of σ. *Parameters:* $a = 1/10$, $\rho = 4\%$, $\pi = 2\%$, $\delta = 1\%$, $\beta = 1.7\%$.

The BGP has a simple structure. We can solve (8) for λ, plug into (7) to obtain σ_c, then plug into (6) to obtain r, and plug into (5) to obtain $\hat{x}$. Finally, once we have the real part of the equilibrium, we use (9) to obtain $\hat{m}$.

The share of wealth in money λ captures the value of liquidity in the economy, and plays a central role. Money provides a safe store of value that improves risk sharing, and it is worth the present value of expenditures on liquidity services. From the definition of h we obtain after some algebra and using the No-Ponzi conditions,[13]

$$m_t + h_t = m_t + \int_t^\infty e^{-r(s-t)} \frac{dM_s}{p_s} = \int_t^\infty e^{-r(s-t)} m_s i ds = \frac{m_t i}{r - (\hat{x} - \delta)} \tag{10}$$

Because of log preferences, we get $m_t i = \rho\beta(k_t + m_t + h_t)$, which yields

$$\lambda \equiv \frac{m_t + h_t}{k_t + m_t + h_t} = \frac{\rho\beta}{r - (\hat{x} - \delta)} \tag{11}$$

Finally, use the Euler equation (5) and the definition of σ_c in (7) to obtain (8).

How big is the value of liquidity λ? In normal times when the real interest rate r is high relative to the growth rate of the economy $\hat{x} - \delta$, the value of liquidity λ is small, close to the expenditure share on liquidity services β. To fix ideas, use a conservative estimate of $\beta = 1.7\%$.[14] But when the real interest rate r is small relative to the growth rate of economy $\hat{x} - \delta$, the value of liquidity can be very large (in the limit $\lambda \to 1$). This happens when idiosyncratic risk σ is large—while

[13]Write $m_t + \int_t^\infty e^{-r(s-t)} \frac{dM_s}{p_s} = m_t + \int_t^\infty e^{-r(s-t)} dm_s + \int_t^\infty e^{-r(s-t)} m_s \pi_s ds = \lim_{T\to\infty} e^{-r(T-t)} m_T + \int_t^\infty e^{-r(s-t)} m_s(r_s + \pi_s) ds$, and use the No-Ponzi condition to eliminate the limit.

[14]As section 2.5 shows, β is the expenditure on liquidity premium across all assets, including deposits and Treasuries. Say checking and savings accounts make up 50% of gdp and have an average liquidity premium of 2%. Krishnamurthy and Vissing-Jorgensen (2012) report expenditure on liquidity provided by Treasuries of 0.25% of gdp. Consumption is 70% of gdp. This yields $\beta = 1.7\%$.

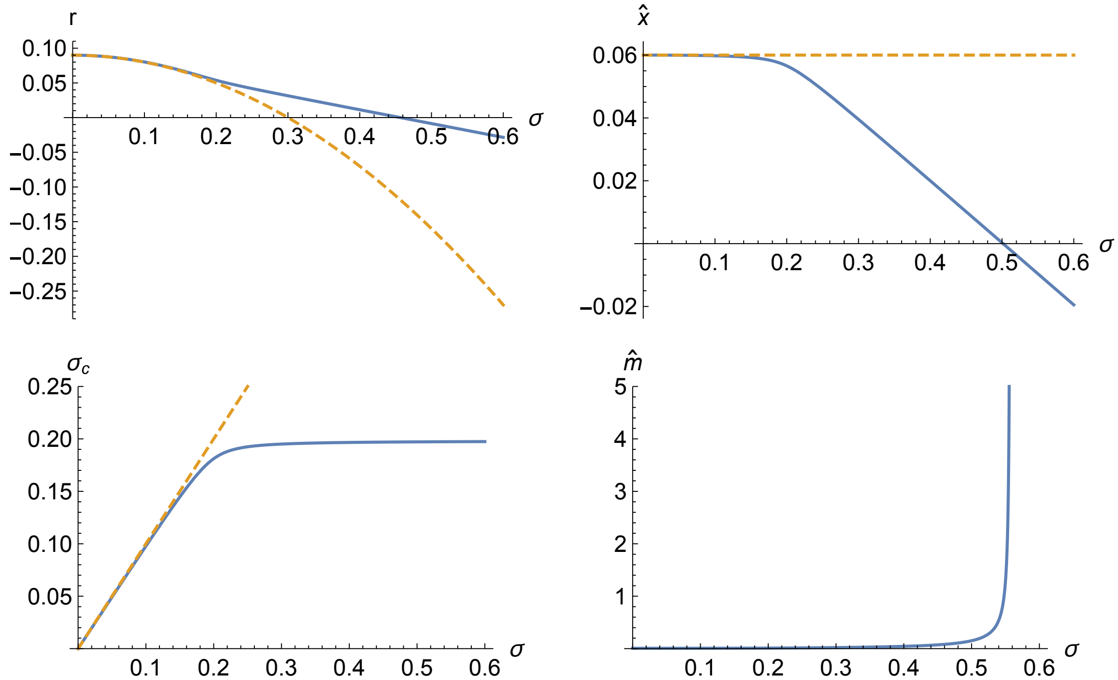

Figure 2: Real interest rate r, investment $\hat{x}$, idiosyncratic consumption risk σ_c, and real money balances $\hat{m}$, as functions of idiosyncratic risk σ, in the non-monetary economy (dashed orange) and the monetary economy (solid blue). *Parameters:* $a = 1/10$, $\rho = 4\%$, $\delta = 1\%$, $\beta = 1.7\%$. Real money balances $\hat{m}$ are computed for a $\pi = 2\%$ inflation target.

capital is discounted with a large risk premium, liquidity is discounted only with the risk-free rate, which must fall when idiosyncratic risk σ is large. Figure 1 shows the non-linear behavior of λ as a function of σ.

It's worth stressing that the value of liquidity includes not only real money balances today m_t, but also future money h_t. In fact, because of log preferences, λ is invariant to the inflation rate π. A higher growth rate of money makes h_t larger, but reduces the real money balances m_t by the same amount.

Proposition 1. *For any $\beta > 0$, the value of liquidity λ is increasing in idiosyncratic risk σ, and ranges from β when $\sigma = 0$ to 1 as $\sigma \to \infty$. Furthermore, idiosyncratic consumption risk $\sigma_c = (1-\lambda)\sigma$ is also increasing in σ, and ranges from 0 when $\sigma = 0$ to $\sqrt{\rho(1-\beta)}$ when $\sigma \to \infty$. For $\beta = 0$, $\lambda = 0$.*

2.3 Non-monetary economy

As a benchmark, consider a non-monetary economy where $\beta = 0$. In this case, $\hat{m} = \hat{h} = 0$ and therefore $\lambda = 0$. The BGP equations simplify to $r = a - \delta - \sigma^2$ and $\hat{x} = a - \rho$.

Higher idiosyncratic risk σ, which makes investment less attractive, is fully absorbed by a lower real interest rate r (and therefore lower nominal interest rate $i = r + \pi$), but a constant investment rate $\hat{x}$ and growth $\hat{x} - \delta$. Figure 2 shows the equilibrium values of r and $\hat{x}$ in a non-monetary economy for different σ (dashed line).

Proposition 2. *Without money ($\beta = 0$), after an increase in idiosyncratic risk σ the real interest rate r falls but investment remains at the first-best level, $\hat{x} = a - \rho$.*

We can understand the response of the non-monetary economy to higher risk σ in terms of the risk premium and the precautionary motive. Use the Euler equation (5) and asset pricing equation (6) to write

$$r = a - \delta - \underbrace{\sigma_c \sigma}_{\text{risk pr.}} \tag{12}$$

$$\hat{x} = \underbrace{a - \rho}_{\text{first-best}} + \underbrace{\sigma_c^2}_{\text{prec. mot.}} - \underbrace{\sigma_c \sigma}_{\text{risk pr.}} = a - \rho \tag{13}$$

Larger risk σ makes capital less attractive, so the risk premium $\alpha = \sigma_c \sigma$ goes up. Other things equal, this depresses investment. But with higher risk the precautionary saving motive σ_c^2 also becomes larger. Agents face more risk and therefore want to save more. Other things equal, this lowers the real interest rate and stimulates investment. Without money $\sigma_c = (1 - \lambda)\sigma = \sigma$, so the precautionary motive and the risk premium cancel each other out and we get the first-best level of investment $\hat{x} = a - \rho$ for any level of idiosyncratic risk σ (this doesn't mean that this level of investment is optimal with $\sigma > 0$).

This is a well known feature of preferences with intertemporal elasticity of one (in the Appendix, I solve the model with general Epstein-Zin preferences).[15] For our purposes, it provides a clean and quantitatively relevant benchmark where higher idiosyncratic risk that makes investment less attractive is completely absorbed by lower real interest rates that completely stabilize investment. But notice in Figure 2 that the real interest rate r could become very negative; in particular, we may need $r \leq \hat{x} - \delta$. This is not a problem without money because capital is risky, but it will be once we introduce money, which is safe, because its value would blow up if $r \leq \hat{x} - \delta$.

2.4 Monetary economy

Now consider the monetary economy with $\beta > 0$, also shown in Figure 2. The presence of money changes how the economy reacts to risk shocks. When σ goes up, money prevents the real interest rate r from falling as much as in the non-monetary economy, and investment $\hat{x}$ falls instead. In particular, without money the real interest rate could be very negative for high σ, but with money it must remain above the growth rate of the economy. These effects are not transitory. It's the new BGP.

To understand the role of money, it's useful to proceed in two steps: (i) money serves as a safe store of value and improves risk sharing, so a large value of liquidity λ keeps the real interest high

[15] Although the real interest rate r always falls with higher risk σ, without money investment $\hat{x}$ may go up or down depending on whether intertemporal elasticity is lower or higher than one. But for relevant parameter values the role of money is the same as in the baseline model with log preferences: it prevents interest rates from falling during downturns and depresses investment relative to the non-monetary economy.

and depresses investment relative to the non-monetary economy; and (ii) the value of liquidity λ endogenously rises during downturns when σ is high.

To understand step (i) use the Euler equation (5), the asset pricing equation (6), and the risk sharing equation (7) to obtain an expression for r and $\hat{x}$ in terms of σ and λ:

$$r = a - \delta - \underbrace{(1-\lambda)\sigma^2}_{\text{risk pr.}} \tag{14}$$

$$\hat{x} = \underbrace{a - \rho}_{\text{first best}} + \underbrace{(1-\lambda)^2\sigma^2}_{\text{prec. mot.}} - \underbrace{(1-\lambda)\sigma^2}_{\text{risk pr.}} = a - \rho - \rho\frac{\lambda - \beta}{1-\lambda} \tag{15}$$

Expressions (14) and (15) show that a larger value of liquidity λ raises the real interest rate and depresses investment. A large value of liquidity λ improves idiosyncratic risk sharing, $\sigma_c = (1-\lambda)\sigma$, which can be seen in the third panel of Figure 2. Essentially, agents with bad shocks sell part of their money holdings to buy more capital and consumption goods from agents with good shocks. As a result, the volatility in their consumption and capital is smaller. Better risk sharing dampens both the risk premium $\sigma_c\sigma$ (raising r) and the precautionary saving motive σ_c^2—but, crucially, it dampens the precautionary motive more. Intuitively, the risk premium comes from the risk of a *marginal* increase in capital holdings, while the precautionary motive comes from the *average* risk in an agent's portfolio, which now includes safe money. Money creates a wedge between the marginal and average risk that weakens the precautionary motive relative to the risk premium. Since the risk premium reduces investment and the precautionary motive increases it, a large value of liquidity λ depresses investment.

To understand step (ii), notice that the value of liquidity λ grows during downturns with high σ, as shown in Figure 1. The value of liquidity λ is equal to the present value of expenditures on liquidity services, as expression (11) indicates. When idiosyncratic risk σ rises, the real interest rate falls relative to the growth rate of the economy because the precautionary motive rises (see the Euler equation (5)), so this present value becomes very large.

Incomplete idiosyncratic risk sharing is essential to the mechanism. If risk sharing is perfect or if there is no idiosyncratic risk, $\sigma = 0$, the monetary economy behaves exactly like the non-monetary one (classical dichotomy). During normal times when idiosyncratic risk σ is small, the role of money is small and can be safely ignored. But it can become very large during periods of high idiosyncratic risk (see Proposition 1).

Proposition 3. *With money ($\beta > 0$) after an increase in idiosyncratic risk σ the real interest rate r falls less than in the economy without money ($\beta = 0$), and investment $\hat{x}$ falls instead, while the value of liquidity λ increases with σ.*

(Classical Dichotomy) If $\sigma = 0$, the real interest rate r and investment $\hat{x}$ are the same in the monetary and non-monetary economies, even though $\lambda = \beta > 0$ in the monetary economy.

It is tempting to interpret the depressed investment as substitution from risky capital to safe money as a savings device; i.e., when capital becomes more risky, it is more attractive to invest in the safe asset. But this is misleading because the economy cannot really invest in money. Goods can be either consumed or accumulated as capital—money is not a substitute for investment in risky capital. What money does is improve how the idiosyncratic risk in capital is shared. Agents with bad shocks use part of their money holdings to buy more capital from those with good shocks.[16] As a result of this risk sharing, the economy substitutes along the consumption-investment margin. To drive home this point, notice that in a model with risky and safe capital (but no money), an increase in risk will typically reduce risky investment but increase the safe one. Money depresses all investment.

Superneutrality and the zero lower bound. While the presence of money has potentially large real effects, money is still neutral and superneutral. Doubling the amount of money would just double prices, leaving all real variables unaffected. With a fixed inflation target π, after an increase in σ real money balances $\hat{m}$ grow as the nominal interest rate $i = r + \pi$ falls. The fourth panel of Figure 2 shows the value of real money balances $\hat{m}$ for an arbitrary inflation target $\pi = 2\%$. A central bank that targets inflation must increase the money supply endogenously after σ increases to keep prices on path. If it didn't, prices would fall but the real allocation wouldn't change.

The inflation target π itself doesn't affect any real variable except real money holdings m. It simply does not appear in equations (14), (15), and (8). As a result, the optimal inflation target is given by the Friedman rule, $i = r + \pi \approx 0$. This maximizes agents' utility from money m without affecting any other real variable. Money superneutrality shows that the mechanism behind the role of money does not hinge on monetary policy. If instead of targeting an inflation rate, the central bank targeted a nominal interest rate i, the behavior of inflation and m would change, but nothing real would be affected. Since $i = r + \pi$, in order to keep the nominal interest rate i constant the central bank would have to raise the inflation target after idiosyncratic risk σ goes up, to compensate for the lower real interest rate r. But this would not affect the real interest rate r or investment $\hat{x}$.

The presence of money does create a zero lower bound on nominal interest rates, $i = r + \pi \geq 0$. But the ZLB does not play any role in mechanism described here. The ZLB places a lower bound on the inflation target π. The central bank is simply unable to deliver an inflation target below $-r$, which does not depend on the inflation target itself. Since r falls with σ, there is a maximum σ such that the equilibrium exists for a given inflation target π. In the fourth panel of Figure 2 we see that, for $\pi = 2\%$, real money balances diverge to infinity as i approaches zero near $\sigma \approx 0.55$. For σ larger than this, the competitive equilibrium is not consistent with a 2% inflation target. But since money is superneutral, changing the inflation target will always "fix" the zero lower bound problem, and it will have no effects on the real allocation or the role of money. In fact, under the optimal monetary policy, $i \approx 0$, the zero lower bound is never a problem.

Proposition 4. *(Superneutrality) With log preferences for liquidity, as long as $\pi \geq -r$, changing the inflation rate π does not affect the value of liquidity λ, the real interest rate r, or investment $\hat{x}$.*

[16] They are not self-insuring in autarky by holding a less risky form of capital. They are sharing idiosyncratic risk.

It only affects real money holdings $\hat{m}$ and the nominal interest rate i.

If the central bank targeted the nominal interest rate $i > 0$, rather than inflation π, the behavior of λ, r, and $\hat{x}$ would not be affected. Only $\hat{m}$, i, and π would be affected.

For any level of idiosyncratic risk σ, the optimal inflation target π delivers the Friedman rule, $i \approx 0$.

It is easy to break superneutrality, but it is a useful theoretical benchmark that highlights that the role of money does not hinge on violating money neutrality and superneutrality. Here superneutrality comes from log preferences, which imply a demand elasticity of money of one. Recall that the value of liquidity is equal to the present value of expenditures on liquidity services $m \times i$. With log preferences a higher nominal interest rate i reduces real money holdings m proportionally, so that $m \times i$ doesn't change. As a result, λ is not affected and neither is any real variable.[17] In section 2.6 I solve the model with a CES demand structure for money and a cash-in-advance constraint. In both cases nominal interest rates have real effects because the expenditure share on liquidity services depends on the nominal interest rate.

2.5 Understanding the mechanism

The presence of money has real effects because it is a safe asset with a liquidity premium. This makes it effectively in positive net supply, which allows it to serve as a safe store of value and improve idiosyncratic risk sharing. Safe assets, private or public, play the same role only to the extent that they have a liquidity premium (such as deposits or short-term Treasuries). That is, to the extent that they are somewhat like money. Private or government safe debt without a liquidity premium does not serve as a safe store of value because it is in zero net supply. I also allow agents to issue outside equity. While equity improves risk sharing, it does not play the sane role as money, because it is also in zero net supply.

To understand the role of the liquidity premium, notice that safe assets without a liquidity premium must be backed by payments with the same present value. Agents may hold the safe assets, but they are also directly or indirectly responsible for the payments backing them. The net value is zero, so they cannot function as a safe store of value. In contrast, assets with a liquidity premium have a value greater than the present value of payments backing them. The difference is the present value of the liquidity premium. This is what makes them a store of value that can improve idiosyncratic risk sharing. Essentially, agents with a bad shock can sell part of their liquid assets to agents with a good shock to reduce the volatility of their consumption. And the net value of these safe liquid assets increases when σ goes up and the real interest rate becomes very low.

It's worth stressing that this is a general equilibrium mechanism. The only reason agents hold money is because it provides liquidity services. From an agent's point of view, risk-free bonds are just as good as a store of value for risk sharing purposes, and they pay interest on top. But agents

[17] This may seem puzzling at first. How can m fall but λ remain constant? Recall that $\lambda = (m + h)/(k + m + h)$ includes not only current real money balances m, but also future money h. As we change the inflation target and i, m and h move in opposite directions.

can't all hold risk-free bonds as a safe store of value. Someone must take the other side and issue risk-free debt. In general equilibrium the real interest rate adjusts to ensure this. Money, and safe liquid assets, have positive net value, so they can improve idiosyncratic risk sharing in general equilibrium.

How does money improve risk sharing? To understand how money improves risk sharing, integrate an individual agent i's dynamic budget constraint (4) to obtain[18]

$$\mathbb{E}^{\tilde{Q}}\left[\int_0^\infty e^{-\int_0^t r_u du}(c_{it} + m_{it}i_t)dt\right] \leq w_0 = k_0 + \int_0^\infty e^{-\int_0^t r_u du} m_t i_t dt \tag{16}$$

Here for simplicity I assume every agent owns an equal part of the aggregate endowment of capital and money. On the left hand side we have the present value of his expenditures on consumption goods and money services. On the right hand side we have the aggregate wealth in the economy, $k_0+m_0+h_0$. The left hand side is evaluated with an equivalent martingale measure $\tilde{Q}$ that captures the market incompleteness; i.e. such that $W_{it} + \int_0^t(\alpha_u/\sigma)du$ is a martingale. A risky consumption plan costs less because it can be dynamically supported with risky investment in capital that yields an excess return α. The endowment of money on the right hand side is safe, however.

With perfect risk sharing, $\sigma = 0$, we have $\alpha_t = 0$, so market clearing $\int m_{it}di = m_t$ means that money drops out of the budget constraint in equilibrium; i.e. $\mathbb{E}^{\tilde{Q}}\left[\int_0^\infty e^{-\int_0^t r_u du} m_{it}i_t dt\right] = \int_0^\infty e^{-\int_0^t r_u du} m_t i_t dt$. Money is worth more than the payments backing it because it has a liquidity premium (that's why it appears on the right hand side), but agents spend on holding money exactly that amount, so it cancels out of the budget constraint and has no effects on the equilibrium.

But if idiosyncratic risk sharing is imperfect, the excess return then is positive, $\alpha_t > 0$. Then even if in equilibrium agents must hold all the money, $\int m_{it}di = m_t$, the present value of expenditures on money services under $\tilde{Q}$ is less than the value of the endowment of money services (which is not risky), $\mathbb{E}^{\tilde{Q}}\left[\int_0^\infty e^{-\int_0^t r_u du} m_{it}i_t dt\right] < \int_0^\infty e^{-\int_0^t r_u du} m_t i_t dt$. As a result, money does not drop out of the budget constraint, and they can use the extra value to reduce the risk in their consumption c_{it}.

To make this clear, agents could choose safe money holdings $m_{it} = m_t$ if they wanted, in which case money would indeed drop out. This corresponds to never trading any money; just holding their endowment. But they are better off trading their money contingent on the realization of their idiosyncratic shocks. They get a risky consumption of money services m_i, but reduce the risk in their consumption c_i. So an agent with a bad idiosyncratic shock in his risky capital can sell part of his money to an agent with a good idiosyncratic shock. Both are better off. The agent with a bad shock gets more consumption and capital than without trading, but less money; the agent with the good shock less consumption and capital, but more money.

[18]The intertemporal budget constraint (16) is equivalent to the dynamic budget constraint (4) with incomplete risk sharing if shorting capital $k_t < 0$ is allowed. This is not required in equilibrium of course.

Government debt, deposits, and Ricardian equivalence. Now let's introduce safe government debt and bank-issued deposits. Both may have a liquidity premium.[19] The bottom line is that government debt and deposits perform the same role as money only if and to the extent that they have a liquidity premium.

Let b_t be the real value of government debt, and $d\tau$ lump-sum taxes. The government's budget constraint is[20]

$$db_t = b_t(i_t^b - \pi)dt - d\tau_t - \frac{dM_t}{p_t}$$

$$dm_t = \frac{dM_t}{p_t} - \pi m_t dt$$

where i_t^b is the nominal interest rate on government bonds. I allow for the possibility that $i_t^b < i_t$ so that government debt also has a liquidity premium. The government has a No-Ponzi constraint $\lim_{T\to\infty} e^{-\int_0^T r_s ds}(b_T + m_T) = 0$. Integrating both equations we obtain

$$m_t + b_t = \int_t^\infty e^{-\int_t^s r_u du}\left(m_s i_s + b_s(i_s - i_s^b)\right)ds + \int_t^\infty e^{-\int_t^s r_u du} d\tau_s \tag{17}$$

The government's total debt is $b_t + m_t$, and it must cover it with the present value of future taxes plus what it will receive because its liabilities b_t and m_t provide liquidity services. When agents hold money, they are effectively paying the government $m_t i_t$ for its liquidity services (the forgone interest); when they hold government debt they are paying $b_t(i_t - i_t^b)$. In particular, if government debt is as liquid as money, $i_t^b = 0$, only the total government liabilities $m_t + b_t$ matter.

There are also banks that can issue deposits d_t that pay interest $i_t^d < i_t$. Banks are owned by households. The net worth of a bank is n_t and follows the dynamic budget constraint

$$dn_t = n_t r_t + d_t(i_t - i_t^d)dt - df_t$$

where f_t are the cumulative dividend payments to shareholders. The bank earns a profit from the spread between the interest it pays on deposits i_t^d and the interest rate at which it can invest, i_t. Using the transversality condition $\lim_{T\to\infty} e^{-rT} n_T = 0$ we can price the bank at v_t:[21]

$$v_t = n_t + \int_t^\infty e^{-\int_t^s r_u du} d_s(i_s - i_s^d)ds$$

The market value of the bank includes its net worth today, plus the present value of profits from the interest rate spread on deposits, $d_t(i_t - i_t^d)$.

Total wealth is $w_t = (k_t - a_t) + d_t + v_t + m_t + (b_t - \int_t^\infty e^{-\int_t^s r_u du} d\tau_s)$, where a_t is the bank's assets. Households own all the capital, money, and government debt (minus the present value

[19] Krishnamurthy and Vissing-Jorgensen (2012) show that US Treasuries have a liquidity premium or convenience yield over equally risk-less private debt. See section 5.

[20] In the baseline model without government debt, we have $b_t = 0$ and $d\tau_t = dM_t/p_t$.

[21] Write $n_t = \int_t^\infty e^{-\int_t^s r_u du} df_s - \int_t^\infty e^{-\int_t^s r_u du} d_s(i_s - i_s^d)ds + \lim_{T\to\infty} e^{-rT} n_T$, and use $v_t = \int_t^\infty e^{-\int_t^s r_u du} df_s$ to obtain $v_t = n_t + \int_t^\infty e^{-\int_t^s r_u du} d_s(i_s - i_s^d)ds$.

of taxes), except for whatever assets the bank holds. They also hold bank debt (deposits) d_t, and bank equity v_t (so they indirectly own the assets that the bank owns). Since the bank's net worth is $n_t = a_t - d_t$, we have $v_t + d_t - a_t = v_t - n_t = \int_t^\infty e^{-\int_t^s r_u du} d_s(i_s - i_s^d) ds$. And $m_t + b_t - \int_t^\infty e^{-\int_t^s r_u du} d\tau_s = \int_t^\infty e^{-\int_t^s r_u du} \left(m_s i_s + b_s(i_s - i_s^b)\right) ds$. So total wealth is

$$w_t = k_t + \int_t^\infty e^{-\int_t^s r_u du} \left(m_s i_s + b_s(i_s - i_s^b) + d_s(i_s - i_s^d)\right) ds \tag{18}$$

Total wealth is capital plus the present value of expenditures on liquidity services, which now include money, liquid government bonds, and deposits, each weighted by its corresponding liquidity premium. Government debt and deposits therefore only have an effect to the extent that they have a liquidity premium. Safe government or private debt without a liquidity premium cancels out and has no effects.

The easiest way to introduce government debt and deposits with a liquidity premium is to put them into the utility function

$$(1 - \beta)\log(c_t) + \beta \log(A(m, b, d))$$

where $A(m, b, d)$ is an homogenous aggregator. Agents will devote a fraction β of expenditures to the liquid aggregate, $\beta = (m_t i + b_t(i - i^b) + d_t(i - i^d))/\text{total expenditures}$. As a result, we just need to reinterpret β as the fraction of expenditures on liquidity services across all assets.[22] The corresponding expression for λ is

$$\lambda = \frac{m_t i + b_t(i - i^b) + d_t(i - i^d)}{r - (\hat{x} - \delta)} \frac{1}{w_t} = \frac{\rho\beta}{r - (\hat{x} - \delta)}$$

In the special case without deposits or government debt we recover expression (11).

Ricardian equivalence holds in this economy. If government debt doesn't have a liquidity premium, changing b_t (and adjusting taxes to service this debt) has no effects on the economy. If government debt has a liquidity premium, then changing b_t can have an effect on the liquidity premium of government debt and perhaps other assets as well. But it will not have any effect on the real side of the economy.

Proposition 5. *(Ricardian Equivalence) With log preferences for liquidity, changes in government debt b have no effects on the real interest rate r, investment $\hat{x}$, or the value of liquidity λ. Changes in b can only affect the liquidity premiums of different assets.*

To see this, notice that the expenditure share on liquidity services across all assets is a constant, β, and this is the only way that liquid government debt can affect the economy. For example, if the liquidity aggregator is Cobb-Douglas, $A(m, b, d) = m^{\epsilon_m} b^{\epsilon_b} d^{\epsilon_d}$ with $\epsilon_m + \epsilon_b + \epsilon_d = 1$, then the

[22] With $i_t - i_t^d > 0$ banks have incentives to supply as much deposits as possible. I'm not providing a theory of what limits them (perhaps capital requirements). But regardless of how we fill in the details of how banks operate, the expenditure share on liquidity services across all assets will be β .

expenditure share on liquidity services from each asset class is fixed; e.g. $b_t(i - i^b)$/expenditures $=$ $\epsilon_b\beta$. Changing b_t only affects the liquidity premium on government bonds, but not on deposits or money.

As with superneutrality, Ricardian equivalence can be broken here if we move away from the log utility over liquidity (see section 2.6 for CES and cash-in-advance formulations). But it's a useful theoretical benchmark that shows that the mechanism behind the role of money does not hinge on a fiscal channel.

Equity markets. The starting point in this paper is that capital is risky and idiosyncratic risk sharing is incomplete. But if agents can hold a diversified (safe) market index, can this function as a safe store of value? Here I'll show that while issuing equity improves risk sharing, it does not perform the same role as money.

In the baseline model agents cannot issue any equity. Let's say instead that they must retain a fraction $\phi \in (0, 1)$ of the equity, and can sell the rest to outside investors. Issuing outside equity improves idiosyncratic risk sharing, of course. Outside investors can fully diversify across all agents' equity, creating a safe market index worth $(1 - \phi)k_t$. If agents could sell all the equity, $\phi = 0$, we would obtain the first best with perfect risk sharing. With $\phi > 0$ we have incomplete idiosyncratic risk sharing.

Since agents can finance an extra unit of capital partly with outside equity, the effective risk of capital for an agent is $\phi\sigma$. In fact, we can obtain the competitive equilibrium by replacing σ with $\phi\sigma$ in (5)-(9). The dynamic budget constraint is now[23]

$$dw_t = (r_t w_t + k_t\alpha_t - c_t - m_t i_t)dt + k_t\phi\sigma dW_t$$

The risk premium is $\alpha_t = \sigma_c(\phi\sigma)$, and the volatility of consumption is $\sigma_c = k_t/(k_t+m_t+h_t)\times(\phi\sigma) = (1 - \lambda)(\phi\sigma)$. The value of liquidity is given by $\lambda = \frac{\rho\beta}{\rho-((1-\lambda)\phi\sigma)^2}$.

But while equity improves risk sharing, it works very differently from money. In particular, without money, $\beta = 0$, an increase in idiosyncratic risk σ is fully absorbed by lower real interest rates $r = a - \delta - (\phi\sigma)^2$, but investment remains at the first best level $\hat{x} = a - \rho$. The reason is that issuing equity improves risk sharing in a way that affects the marginal risk from an extra unit of capital and the average risk in agent's portfolio equally. As a result, it dampens the risk premium $\sigma_c\phi\sigma = (\phi\sigma)^2$ and the precautionary motives $\sigma_c^2 = (\phi\sigma)^2$ equally, canceling out. And the value of equity is backed by the firm's assets, so it's not a positive net value. The aggregate wealth in the economy is still given by the right hand side of (16), but the total value of capital is split into inside and outside equity $k_t = \phi k_t + (1 - \phi)k_t$.[24] In particular, the value of the market index does not

[23] Equity can be diversified so its return must be r. In equilibrium agents are holding $w_t = n_t + m_t + h_t + e_t$ where $n_t = \phi k_t$ is the inside equity in their firm that they retain, and $e_t = (1 - \phi)k_t$ is the diversified outside equity in other agents' firms. Total equity is worth $n_t + e_t = k_t$; since there are no adjustment costs, Tobin's q is 1 here. Both inside and outside equity yield r, but the inside equity has idiosyncratic risk (outside equity also has idiosyncratic risk, but it gets diversified). Agents therefore also get a wage or bonus as CEO of their firm to compensate them for the undiversified idiosyncratic risk, $k_t\alpha_t$.

[24] More generally, if firms use debt, $k_t = n_t + e_t + d_t$, where n_t is inside equity, e_t is outside equity, and d_t is debt.

blow up to infinity as r approaches the growth rate $\hat{x} - \delta$, as the value of liquidity does.[25]

An analogy with an infinitely-lived safe tree. The main assumption in this paper is that real investments are risky and this risk cannot be fully shared. But to understand the role of money as a store of value, it is useful to study what would happen if there was a safe tree. There are similarities and differences with how money works in the model.

Let's introduce an infinitely-lived safe tree as close as possible to money. Suppose the economy has a tree that produces a safe flow of fruit (apples), that enters the utility function analogously to money,

$$\mathbb{E}\left[\int_0^\infty e^{-\rho t}\left((1-\beta)\log c_t + \beta \log a_t\right) dt\right]$$

where c_t represents the consumption goods produced by (risky) capital, and a_t represents apples produced by the tree. The tree cannot be produced and apples cannot be used to produce capital. The tree does not enter the resource constraint for goods in any way, just like money.

Households will devote a fraction β of their expenditures to apples, $p_{at}a_t$, and the value of the tree will be

$$q_t = \int_t^\infty e^{-\int_t^s r_u du} p_{as} a_s ds$$

This is analogous to expression (10) for the value of money. Total wealth in the economy will therefore be $w_t = k_t + q_t$. In a BGP, the value of the tree q_t grows at the same rate as capital. Slightly abusing notation, let $\lambda = q_t/w_t$,[26]

$$\lambda = \frac{\rho\beta}{r - (\hat{x} - \delta)}$$

which is analogous to expression (11) for λ.

Idiosyncratic risk in consumption will be $\sigma_c = (1-\lambda)\sigma$, so the model will behave exactly like the baseline model with money. The safe tree has positive net value, so it will improve risk sharing and keep the real interest rate r high and investment $\hat{x}$ depressed relative to the economy without the apple tree. And an increase in idiosyncratic risk σ will increase the value of the apple tree, increasing the gap between the economy with and without the tree.

Of course, if the safe tree can be produced, an increase in idiosyncratic risk of capital σ will reduce investment in risky capital but increase investment in the safe tree. In contrast, money and other safe assets with a liquidity premium depress all investment. More importantly, the starting point of this paper is that real investment—capital, land, housing—is risky and risk cannot be perfectly shared. A possible exception is something like gold, which is safe and durable.

All the financial claims on firms add up to the value of their assets.

[25] Total equity is always worth total capital, whose price takes into account its uninsurable idiosyncratic risk. As σ grows and r drops, insider wages or bonuses αk_t increase to compensate for the idiosyncratic risk.

[26] Write $q_t/w_t = (1/w_t)\int_t^\infty e^{-r(s-t)}\rho\beta w_s ds = (\rho\beta w_t/w_t)\int_t^\infty e^{-r(s-t)}e^{(\hat{x}-\delta)(s-t)}ds$.

Wedge accounting with a consumption tax. We can further understand the role of money by asking what kind of distortionary taxes, or wedges, can reproduce its equilibrium real effects. To this end, consider an economy without money, $\beta = 0$, but with a consumption tax τ_c. For now, let the tax revenue be rebated immediately through lump-sum transfers $\tau_t = \tau_c c_t$, so the government runs a balanced budget. The consumption tax does not distort the Euler equation (5) or the asset pricing equation for capital (6), but the expression for idiosyncratic risk must now be modified to

$$\sigma_c = \frac{k_t}{k_t + z_t}\sigma$$

where $z_t = \int_t^\infty e^{-\int_t^s r_u ds} \tau_c c_s ds$ is the present value of the consumption tax. In a BGP, $z_t = \frac{\tau_c c_t}{r - (\hat{x} - \delta)}$, and we know that consumption is given by $c_t(1 + \tau_c) = \rho(k_t + z_t)$ because of log preferences. Slightly abusing notation, let $\lambda = \frac{z_t}{k_t + z_t}$ and, following the same steps as in the baseline, we obtain $\sigma_c = (1 - \lambda)\sigma$ and

$$\frac{z_t}{k_t + z_t} \equiv \lambda = \frac{\rho \frac{\tau_c}{1+\tau_c}}{\rho - ((1 - \lambda)\sigma)^2}$$

If we set $\tau_c = \frac{\beta}{1-\beta}$, we obtain the same allocation for consumption and capital as in the economy with money. In other words, the equilibrium effects of money are equivalent to introducing a distortionary consumption tax. The consumption tax provides risk sharing because the tax paid by each agent depends on their idiosyncratic history. Agents with a string of good shocks will have higher consumption and therefore pay more in taxes. In the case of money, agents with good shocks willingly spend more on liquidity services.

It's worth stressing that Ricardian equivalence still holds here. If the government had some debt b_t instead of running a balanced budget, it wouldn't have any real effects as long as we don't change the distortionary consumption tax (that is, if we financed the extra debt with non-distortionary lump-sum taxes). Alternatively, if the government lowers the consumption tax, it will have real effects regardless of whether it finances the reduction with debt or with higher lump-sum taxes. It's not government debt that has real effects (unless it has a liquidity premium), but rather the distortionary tax.

2.6 Cashless limit and alternative specifications of money demand

In this section I show that the real effects of money survive in the cashless limit where $\beta \to 0$. In addition, I consider alternative specifications for money demand.

Cashless limit. As explained in section 2.2, the value of money λ is the present value of expenditures on liquidity discounted at the risk-free rate. When the real interest rate is high relative to the growth rate of the economy, λ is small; close to the expenditure share on liquidity services β. But when the real interest rate is very close to the growth rate of the economy, λ can become very large regardless of how small β is.

So if we take the cashless limit, $\beta \to 0$, the competitive equilibrium will not always converge

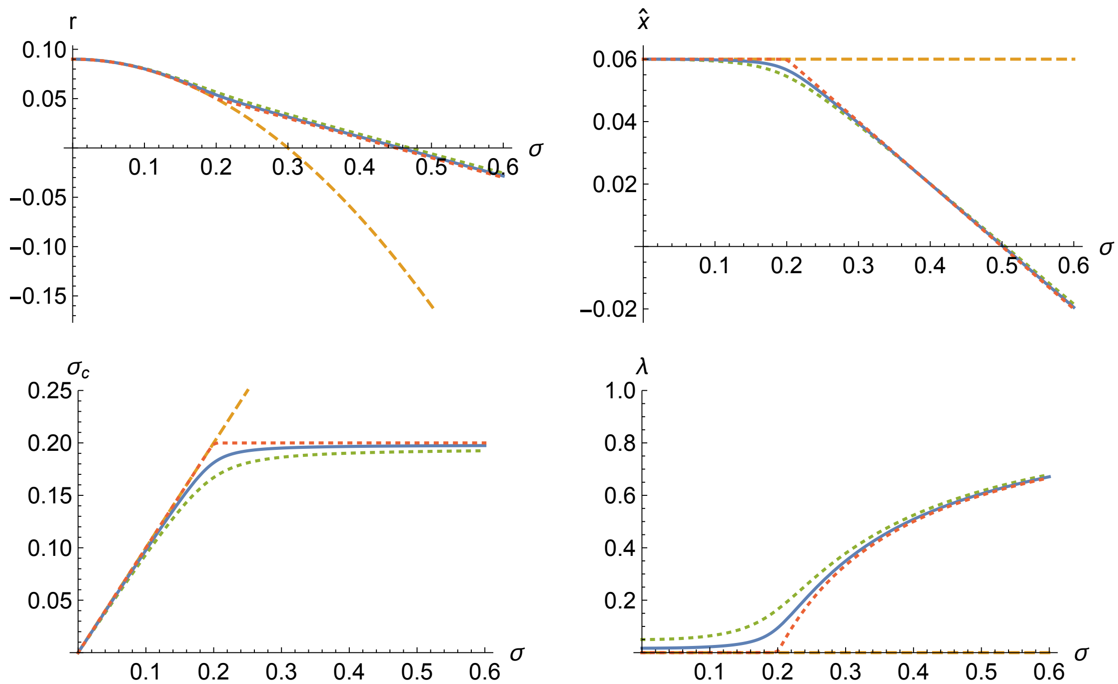

Figure 3: (Cashless limit) The real interest rate r, investment $\hat{x}$, idiosyncratic consumption risk σ_c, and value of liquidity λ as function of σ, for $\beta = 5\%$ (dotted green), $\beta = 1.7\%$ (solid blue—baseline case), $\beta = 0.01\%$ (dotted red), and $\beta = 0$ (dashed orange—non-monetary economy). *Other parameters: $a = 1/10$, $\rho = 4\%$, $\delta = 1\%$.*

to that of the non-monetary economy with $\beta = 0$. For σ such that in the non-monetary economy the real interest rate is above the growth rate, the monetary economy will indeed converge to the non-monetary one as $\beta \to 0$. But for σ such that in the non-monetary economy the real interest rate is equal or below the growth rate of the economy, this cannot happen. As the real interest rate drops and approaches the growth rate of the economy, the value of liquidity λ blows up to keep r above $\hat{x} - \delta$, no matter how small β is. As a result, the real effects of money survive even in the cashless limit $\beta \to 0$, with high interest rates and depressed investment relative to the non-monetary economy. Figure 3 shows the convergence to the cashless limit.

Proposition 6. *If $\sigma < \sqrt{\rho}$ then as $\beta \to 0$ the competitive equilibrium converges to that of a non-monetary economy with $\beta = 0$. But if $\sigma \geq \sqrt{\rho}$, it converges to an equilibrium where the real interest rate is high and investment low relative to the non-monetary economy with $\beta = 0$.*

It is important to make sure we are not violating any Ponzi conditions. Proposition 1 ensures that $\sigma_c^2 = ((1-\lambda)\sigma)^2 < \rho$ for all σ and any $\beta > 0$, so the Euler equation (5) guarantees that $r > \hat{x} - \delta$. But what happens if $\beta = 0$? Then the only value of λ that satisfies the No-Ponzi condition is $\lambda = 0$. If $\sigma \geq \sqrt{\rho}$ the limit of the monetary equilibrium as $\beta \to 0$ would be an equilibrium of the non-monetary economy with $\beta = 0$ except for the No-Ponzi conditions. In other words, the monetary economy, which cannot have bubbles, converges to a bubbly equilibrium of the non-monetary economy. I will discuss the link with bubbles in detail in section 5.

Alternative specifications of money demand: CES and cash-in-advance. The baseline specification with log preferences has a constant expenditure share on liquidity services β. Here I show that the role of money is robust to two alternative specifications: CES preferences over consumption and money, and a cash-in-advance constraint (CIA). As in the baseline setting, with perfect risk sharing the presence of money has no effect on the equilibrium, and money is superneutral. With incomplete risk sharing, however, the presence of money has real effects. The only modification in the model is that the expenditure share on liquidity services $\tilde{\beta}(i)$ becomes a function of the nominal interest rate. As a result, money is not superneutral and monetary policy becomes important.

First, consider a CES specification for money demand

$$\mathbb{E}\left[\int_0^\infty e^{-\rho t} \log\Big(\Big((1-\beta)^{\frac{1}{\eta}} c_t^{\frac{\eta-1}{\eta}} + \beta^{\frac{1}{\eta}} m_t^{\frac{\eta-1}{\eta}}\Big)^{\frac{\eta}{\eta-1}}\Big) dt\right]$$

where $\eta < 1$ is the demand elasticity of money. With $\eta = 1$ we recover the baseline setting.[27]

With these preferences, the expenditure share on liquidity services is a function of the nominal interest rate,

$$\frac{mi}{\rho w} = \tilde{\beta}(i) = \frac{\beta i^{1-\eta}}{1-\beta+\beta i^{1-\eta}}$$

The expenditure share $\tilde{\beta}(i)$ is always positive and it's increasing in i because $\eta < 1$.

The Euler equation and asset pricing equation for capital are unchanged. Consumption and money demand are still proportional to wealth, and share the same idiosyncratic risk σ_c. The expressions for r, $\hat{x}$, and λ are the same as in the baseline setting, except that we need to replace β with $\tilde{\beta}(i) = \tilde{\beta}(r+\pi)$. As a result, the presence of money still has real effects. For a given $\sigma > 0$, the monetary economy has a higher real interest rate r and lower investment $\hat{x}$, compared to the non-monetary economy.

The main difference with the baseline setting is that money is not superneutral when risk sharing is incomplete. Changes in the inflation target have real effects because they change the expenditure share on liquidity services, $\tilde{\beta}(r+\pi)$, and therefore the value of liquidity,

$$\lambda(r+\pi) = \frac{\rho\tilde{\beta}(r+\pi)}{\rho - ((1-\lambda(r+\pi))\sigma)^2}$$

Furthermore, the monetary policy rule can affect how the economy responds to an increase in risk σ, through the endogenous response of monetary policy. A fixed inflation target π implies that as the real interest rate r drops, so does the nominal interest rate $i = r + \pi$, which affects $\tilde{\beta}(i)$. If instead we follow a nominal interest rate target i (adjusting the inflation target appropriately), the expenditure share on liquidity services $\tilde{\beta}(i)$ remains unchanged in response to the increase in σ, and the equilibrium response of the real interest rate r and investment $\hat{x}$ is therefore the same as in the baseline with log preferences.

[27] Everything goes through with $\eta > 1$, but $\eta < 1$ is the empirically relevant case.

The CIA case is similar. Now agents have log utility only over consumption, but must respect a CIA constraint $c_t \leq vm_t$, where v is money velocity. The CIA constraint does not affect the Euler equation or the asset pricing equation, but the expenditure share on liquidity services is now increasing in the nominal interest rate,

$$\tilde{\beta}(i) = \frac{i/v}{1 + i/v} \geq 0$$

The presence of money still has the same real effects as in the baseline case and, just as with CES preferences, money is not superneutral. Changing the inflation target has real effects through $\tilde{\beta}(i)$, and the monetary policy rule affects how the economy responds to an increase in idiosyncratic risk σ.

It's worth highlighting the special case where $i \to 0$ and therefore $\tilde{\beta}(i) \to 0$. Both with CES preferences and with the CIA constraint, the real effects of money do not disappear, essentially for the same reason that they don't disappear in the cashless limit with $\beta \to 0$. For $\sigma > \sqrt{\rho}$, while the expenditure share on liquidity services converges to zero, $\tilde{\beta}(i) \to 0$, the value of liquidity $\lambda \to 1 - \sqrt{\rho}/\sigma > 0$ does not converge to zero because the real interest rate approaches the growth rate of the economy.[28]

3 Efficiency

In this section I study the efficiency properties of the monetary competitive equilibrium. Money provides a safe store of value that prevents the real interest rate from falling and depresses investment. This is costly because we get low investment, but in exchange we get better idiosyncratic risk sharing. The main result in this section is that the monetary competitive equilibrium is inefficient. When idiosyncratic risk σ is low, there is too little risk sharing and investment is too high. But when idiosyncratic risk is high, there is too much risk sharing and investment is too low.

I first microfound the reduced-form incomplete risk sharing constraint with a moral hazard problem with hidden trade, so that the competitive equilibrium studied in section 2 is the result of allowing agents to write privately optimal contracts. I then characterize the optimal allocation in this environment.

Ultimately, the inefficiency comes from the presence of hidden trade in the environment.[29] But implementing the optimal allocation does not involve monetary policy (recall that changing inflation targets has no real effects). The optimal allocation can be implemented with a tax or subsidy to capital, which internalizes the externality.

[28] In the CIA case we can actually set $i = 0$, but doing this requires $r = \hat{x} - \delta$ if $\sigma > \sqrt{\rho}$, so we get a bubble. This is not surprising, since as $\beta \to 0$ the monetary economy approaches the non-monetary economy with a bubble. See section 5 for a discussion of the link with bubbles.

[29] See Kehoe and Levine (1993) and Farhi et al. (2009). Di Tella (2016) has a similar contractual setting with hidden trade but without hidden savings. Instead, there is an endogenous price of capital. There is an externality because the private benefit of the hidden action depends on the value of assets. This is absent here because the equilibrium price of capital is always one. But the externality here, produced by hidden intertemporal trade, is absent from that paper.

3.1 Setting

I provide the microfoundations for the reduced-form incomplete idiosyncratic risk sharing assumed in the baseline model in a setting with moral hazard and hidden trade.[30] See the Online Appendix for technical details.

Agents can write complete, long-term contracts with full commitment. A contract $\mathcal{C} = (c, m, k)$ specifies how much the agent should consume c_t, hold money m_t, and capital k_t, as functions of his report of his own idiosyncratic shock $Y_t = W_t - \int_0^t \frac{s_u}{\sigma} du$. The problem is that the shock W_t itself is not observable, so the agent can misreport at rate s_t. If the principal sees low returns reported, he doesn't know if the true returns were low or the agent was misreporting.

Misreporting allows the agent to divert returns to a private account. Importantly, the agent doesn't have to immediately consume what he steals. He has access to hidden trade that allows him to choose his actual consumption $\tilde{c}$, money $\tilde{m}$, and capital $\tilde{k}$. His hidden savings n satisfy a dynamic budget constraint

$$dn_t = (n_t r_t + c_t - \tilde{c}_t + (m_t - \tilde{m}_t)i_t + (\tilde{k}_t - k_t)\alpha_t + k_t s_t)dt + (\tilde{k}_t - k_t)\sigma dW_t \tag{19}$$

with solvency constraint $n_t \geq \underline{n}_t$, where $\underline{n}_t$ is the natural debt limit.[31] It is without loss of generality to implement no stealing and hidden trades in the optimal contract. A contract is *incentive compatible* if

$$(c, m, k, 0) \in \arg \max_{(\tilde{c}, \tilde{m}, \tilde{k}, s)} U(\tilde{c}, \tilde{m}) \quad st: \quad (19) \tag{20}$$

An incentive compatible-contract is *optimal* if it minimizes the cost of delivering utility to the agent:

$$J(u_0) = \min_{(c,m,k) \in IC} \mathbb{E}\left[\int_0^\infty e^{-rt}(c_t + m_t i_t - k_t \alpha_t)dt\right] \quad st: \quad U(c, m) \geq u_0 \tag{21}$$

In general this could be a difficult problem to solve, because the hidden trade gives the agent a very rich set of deviations. However, in this case the optimal contract can be characterized in a straightforward way, as the solution to the portfolio problem in section 2. We say that contract (c, m, k) *solves the portfolio problem for* $w_0 > 0$ if it maximizes $U(c, m)$ subject to the dynamic budget constraint (4).

Proposition 7. *Let (c, m, k) be an optimal contract for initial utility u_0, with cost $J(u_0)$. Then (c, m, k) solves the portfolio problem for $w_0 = J(u_0)$.*

Conversely, let (c, m, k) solve the portfolio problem for some $w_0 > 0$. If in addition $\lim_{t\to\infty} \mathbb{E}[e^{-rt} w_t] = 0$,[32] *then (c, m, k) is an optimal contract for initial utility u_0 with $J(u_0) = w_0$.*

Proof. See Online Appendix. □

[30] The environment is based on Di Tella and Sannikov (2016).

[31] The natural debt limit is $\underline{n}_t = \max_s \mathbb{E}^{\tilde{Q}}\left[\int_t^\infty e^{-\int_t^u}(c_u(Y^s) + m_u(Y^s)i_u + k_u(Y^s)s_u)du\right]$. This is the maximum amount the agent can pay back for sure. See Online Appendix for details.

[32] This condition will always be satisfied in the monetary equilibrium.

Proposition 7 means that the competitive equilibrium characterized in section 2 can also be interpreted as the outcome of allowing agents to write privately optimal contracts in this environment. The intuition is easy to grasp. The principal can consume, save, and invest on his own, so the principal essentially has no tools he can use to discipline the agent, and can only give him risk-free debt. Under those conditions, the optimal contract is implemented by letting the agent choose his consumption-portfolio plan on his own. This also ensures global incentive compatibility.

To understand this environment, write the local incentive compatibility constraints,[33]

$$\sigma_{ct} = \rho(1-\beta)c_t^{-1}k_t\sigma \qquad \text{"skin in the game"} \tag{22}$$

$$\mu_{ct} = r_t - \rho + \sigma_{ct}^2 \qquad \text{Euler equation} \tag{23}$$

$$\alpha_t = \sigma_{ct}\sigma \qquad \text{demand for capital} \tag{24}$$

$$m_t/c_t = \beta/(1-\beta)i_t^{-1} \qquad \text{demand for money} \tag{25}$$

The "skin in the game" constraint (22) says that the agent must be exposed to his own idiosyncratic risk to align incentives. The agent could always misreport a lower return and consume those funds, so incentive compatibility requires that the present value of his consumption goes down by $k_t\sigma$ after bad reported outcomes Y_t. The skin in the game constraint is expressed in terms of the volatility of his consumption σ_{ct}. If he steals a dollar, he won't consume the dollar right away; he will consume it only at rate $\rho(1-\beta)$.[34] So his consumption must be exposed to his idiosyncratic shock as in (22). This is costly, of course. In the first best we would have perfect idiosyncratic risk sharing, $\sigma_{ct} = 0$, but we need to expose the agent to risk to align incentives.

The other IC constraints (23), (24), and (25) come from the agent's ability to save at the risk-free rate, secretly invest in capital, and choose his money holdings, respectively. Ultimately they arise from agents' ability to secretly trade amongst themselves. These constraints are binding. The principal would like to follow the Inverse Euler equation, and to front-load the agent's consumption to improve risk sharing.

The tradeoff between intertemporal consumption smoothing and idiosyncratic risk sharing captured in the skin in the game constraint (22) is central to the equilibrium effects of money. First, we'd like to see how this constraint manifests in the competitive equilibrium. Write $\sigma_{ct} = (1-\lambda)\sigma = (k_t/w_t)\sigma$. Using $c_t = \rho(1-\beta)w_t$, we obtain equation (22). Using the resource constraint, we obtain

$$\sigma_{ct} = \frac{\rho(1-\beta)}{a-\hat{x}_t}\sigma \tag{26}$$

The IC constraint links idiosyncratic risk sharing and investment. Higher investment $\hat{x}_t$ requires

[33] The competitive equilibrium and the planner's allocation will be BGPs, but it is important to allow for time-varying allocations and prices.

[34] An equivalent derivation: the agent's continuation utility if he doesn't misbehave, U_t, follows a promise-keeping constraint $dU_t = \left(\rho U_t - \left(\beta \log(c_t) + (1-\beta)\log(m_t)\right)\right) dt + \sigma_{Ut} dW_t$. If he misreports he can immediately consume what he stole (he is indifferent at the margin) and obtain utility $(1-\beta)c_t^{-1}k_t$, so incentive compatibility requires $\sigma_{Ut} = (1-\beta)c_t^{-1}k_t\sigma$. Because the agent can secretly save and invest, his continuation utility must be $U_t = A + \frac{1}{\rho}\log(c_t)$, so we get $\sigma_{ct} = \rho\sigma_{Ut}$.

exposing agents to more idiosyncratic risk σ_{ct}. When the value of liquidity λ goes up and improves risk sharing, it is moving the equilibrium along this IC constraint. In equilibrium this must be consistent with individual optimization, captured by the risk premium and the precautionary motive. As we'll see, the planner will choose a different point on this IC constraint.

All these conditions are only necessary, and are derived from considering local, single deviations by the agent. Establishing global incentive compatibility is difficult in general, but in this environment it's straightforward. Because the optimal contract coincides with the optimal portfolio problem where the agent essentially does what he wants, global incentive compatibility is ensured.

3.2 Planner's problem

The planner faces the same environment with moral hazard and hidden trade.[35] An *allocation* is a plan for each agent (c_i, m_i, k_i) and aggregate consumption c, investment x, and capital k satisfying the resource constraints (1), (2), $c_t = \int_0^1 c_{i,t} di$ and $k_t = \int_0^1 k_{i,t} di$. An allocation is *incentive compatible* if there exist processes for real interest rate r, nominal interest rate i, and idiosyncratic risk premium α, such that (20) holds for each agent.

The local IC constraints are necessary for an incentive-compatible allocation. But notice that constraints (23), (24), and (25) involve prices that the planner doesn't take as given. What these constraints really say is that all agents must be treated the same, or else they would engage in hidden trades amongst themselves. So all agents get the same μ_c, σ_c, m/c, and k/c, and only differ in the scale of their contract, reflecting their initial utility and their idiosyncratic history. But the planner already wants to treat all agents the same, so these constraints are not binding for the planner. This is why the planner can improve over the competitive equilibrium.

The only true constraint for the planner is the skin in the game constraint (22), which can be re-written using the resource constraints as (26). Both the competitive equilibrium and the planner must respect the link between investment and idiosyncratic risk sharing induced by incentive compatibility.

Using the aggregate resource constraints (1) and (2), and the incentive compatibility constraints, we can write the planner's objective function[36]

$$\mathbb{E}\left[\int_0^\infty e^{-\rho t}\left((1-\beta)\log(c_{i,t}) + \beta\log(m_{i,t})\right)dt\right]$$

[35] It is natural to wonder if the planner could simply refuse to enforce debt contracts in order to eliminate hidden trade. Here I'm assuming the hidden trade is a feature of the environment that the planner cannot change; e.g. agents may have a private way of enforcing debt contracts. As we'll see, the hidden trade constraints are already not binding for the planner, so he wouldn't gain anything from doing this. And we wouldn't learn a lot from pointing out that the planner could do better if he can change the environment.

[36] First compute $\mathbb{E}[\log c_{it}]$ and $\mathbb{E}[\log m_{it}]$ to obtain

$$= \mathbb{E}\left[\int_0^\infty e^{-\rho t}\left((1-\beta)\log(c_t) + \beta\log(m_t) - \frac{1}{2}\frac{\sigma_{ct}^2}{\rho}\right)dt\right]$$

Then write $c_t = k_t(a - \hat{x}_t)$, $m_t = \frac{\beta}{1-\beta}\frac{c_t}{i_t}$, and compute $\mathbb{E}[\log k_t]$ to obtain the final expression.

$$= \mathbb{E}\left[\int_0^\infty e^{-\rho t}\left(\log(k_0) + \log(a - \hat{x}_t) + \frac{\hat{x}_t - \delta}{\rho} - \frac{1}{2}\frac{\sigma_{ct}^2}{\rho} + \beta \log\left(\frac{\beta}{1-\beta} i_t^{-1}\right)\right) dt\right] \tag{27}$$

The planner's problem then boils down to choosing processes for the aggregate consumption c, investment x, and nominal interest rate i. Regardless of the allocation for consumption and capital, the optimal monetary policy is to follow the Friedman rule, $i \approx 0$, and deliver unbounded utility from real money balances.[37] So we proceed in two steps. First, for an arbitrary path for nominal interest rates i, we say the allocation of consumption and capital is *optimal* if it maximizes (27) subject to (26). We are trading off consumption, investment, and risk sharing, taking as given the monetary policy. The solution to this problem is independent of the path for nominal interest rates i. Then, given the optimal allocation of consumption and capital, we implement the Friedman rule.

The FOC for $\hat{x}_t$ yields

$$\hat{x} = \underbrace{a - \rho}_{\text{first-best}} - \sigma_c^2 \tag{28}$$

where recall that $\sigma_c = \frac{\rho(1-\beta)\sigma}{a-\hat{x}}$. The first term is the first-best level of investment, $a - \rho$, corresponding to perfect risk sharing. The second term reflects that higher investment tightens the IC constraint (22) and increases agents' exposure to idiosyncratic risk. The planner realizes that he can improve risk sharing if he is willing to distort the intertemporal consumption margin. Private contracts also realize this, but they are constrained by agents' access to hidden trade. They must respect agents' Euler equation (23) and demand for capital (24) taking r and $\alpha = a - \delta - r$ as given.

The planner reduces investment $\hat{x}$ to improve idiosyncratic risk sharing σ_c. This tradeoff is more attractive when idiosyncratic risk σ is higher. So investment $\hat{x}$ falls with σ, but idiosyncratic consumption risk σ_c goes up less than proportionally to σ. In the background, the real interest rate r falls with σ.

Proposition 8. *In the planner's optimal allocation, an increase in idiosyncratic risk σ depresses investment $\hat{x}$ and the real interest rate r, and increases idiosyncratic consumption risk σ_c, but less than proportionally, i.e. σ_c/σ falls. When $\sigma = 0$ we have the first best investment and risk sharing, with $\sigma_c = 0$, $\hat{x} = a - \rho$, and $r = a - \delta$.*

3.3 Competitive equilibrium vs. planner's allocation

In the competitive equilibrium, money provides a safe store of value that improves risk sharing but depresses investment. This is the same tradeoff that the planner considers, but the competitive equilibrium doesn't do it efficiently. When idiosyncratic risk σ is low, money provides too little insurance and investment $\hat{x}$ is too high. When idiosyncratic risk is large, money provides too much insurance and investment is too low.

Figure 4 compares the competitive equilibrium and the planner's allocation. Investment $\hat{x}$ and consumption risk σ_c are too high in the competitive equilibrium relative to the planner's allocation

[37] Because of the log preferences we can't set $i = 0$ because we would get infinite utility. But $i = 0$ is optimal in a limiting sense.

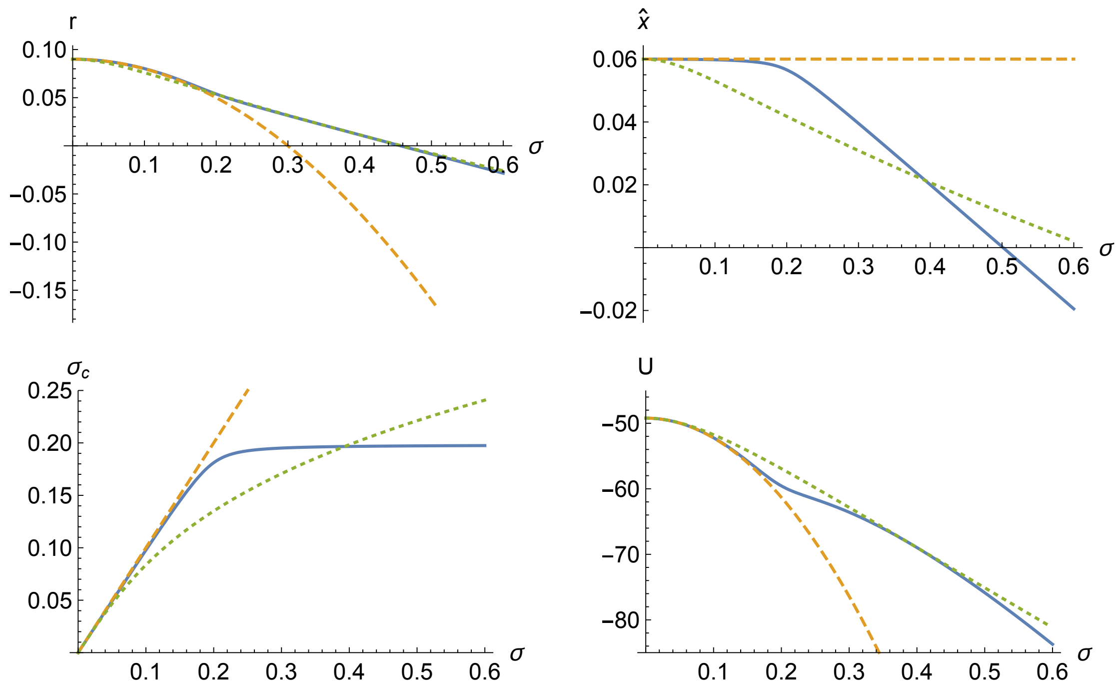

Figure 4: Interest rate r, investment $\hat{x}$, idiosyncratic risk σ_c, and utility from consumption in the non-monetary economy (dashed orange), monetary competitive equilibrium (solid blue), and the social planner's allocation (dotted green). *Parameters:* $a = 1/10$, $\rho = 4\%$, $\delta = 1\%$, $\beta = 1.7\%$.

when σ is low, and too low when σ is high. In contrast, in the non-monetary economy, investment $\hat{x}$ and consumption risk σ_c are always too high.

The fourth panel shows the utility under each allocation. To make utility comparable with the non-monetary economy, I set $\log(\beta i^{-1}/(1-\beta)) = 0$.[38] As expected, the planner's allocation always delivers more utility than the competitive equilibrium. But the monetary competitive equilibrium delivers more utility than the non-monetary allocation. Money depresses investment relative to the first-best, but it provides valuable risk sharing.

Figure 4 also shows the behavior of the equilibrium interest rate r in the competitive equilibrium and the planner's allocation. Perhaps surprisingly, the real interest rate is very close in both allocations, and higher than in the non-monetary economy. There are two conflicting forces. Lower investment $\hat{x}$ lowers the growth rate of consumption and therefore r. But lower idiosyncratic risk σ_c dampens the precautionary saving motive, raising r.

To understand the difference between the competitive equilibrium and the planner's allocation, we can compare the planner's FOC for $\hat{x}$, given by (28), with investment in the competitive equi-

[38]Strictly speaking, the monetary and non-monetary economy are different environments, so welfare cannot be compared in a straightforward way except in the cashless limit. This condition ensures that $(1-\beta)\log c + \beta \log m = \log c$, so we can compare utility in the monetary and non-monetary economies. In the cashless limit, $\beta \to 0$, this is true for any $i > 0$.

librium

$$\hat{x} = a - \rho - \sigma^2 \times (\sigma_c/\sigma)^2 \qquad \text{social planner} \tag{29}$$

$$\hat{x} = a - \rho - \sigma^2 \times (\sigma_c/\sigma) \times (1 - \sigma_c/\sigma) \qquad \text{competitive equilibrium} \tag{30}$$

The competitive equilibrium and the planner share the same expression for incentive compatibility

$$\sigma_c/\sigma = \frac{\rho(1-\beta)}{a - \hat{x}} \tag{31}$$

This pins down the set of $\hat{x}$ and σ_c/σ that are incentive compatible, but the planner and the competitive equilibrium disagree on which $(\hat{x}, \sigma_c/\sigma)$ pair to pick. We can interpret (29) and (30) as the desired investment $\hat{x}$ for a given σ_c. The intersection with (31) gives the planner solution or the competitive equilibrium, respectively.

Figure 5 captures the situation.[39] Investment $\hat{x}$ for the competitive equilibrium and the planner's allocation intersect at $\sigma_c/\sigma = \frac{1}{2}$. If they cross the IC constraint below that level, investment is too low in the competitive equilibrium, and too high if they cross it above that level. Which is the case depends on the level of idiosyncratic risk σ, and how much risk sharing money provides. If $\beta \geq 1/2$ the value of liquidity is very large, so there is too much risk sharing and too little investment in the competitive equilibrium for any σ. For the quantitatively relevant case with $\beta < 1/2$, the value of liquidity is too small for low σ, so there is too little risk sharing and too much investment. But for high σ the value of liquidity is too large, so there is too much risk sharing and too little investment.

Proposition 9. *If $\beta \in (0, 1/2)$, there is a $\sigma^* = 2\sqrt{\rho(1-2\beta)} > 0$ such that for $\sigma \in (0, \sigma^*)$ investment and consumption risk are too high in the competitive equilibrium, compared to the planner's allocation; that is, $\hat{x}^{CE} > \hat{x}^{SP}$ and $\sigma_c^{CE} > \sigma_c^{SP}$. For $\sigma > \sigma^*$ investment and consumption risk are too low in the competitive equilibrium; that is, $\hat{x}^{CE} < \hat{x}^{SP}$ and $\sigma_c^{CE} < \sigma_c^{SP}$. If $\beta \in [1/2, 1)$ investment and consumption risk are too low in the competitive equilibrium for any $\sigma > 0$.*

3.4 Understanding the inefficiency

To understand the inefficiency in this economy, it's useful to focus on the cashless limit, $\beta \to 0$, and compare the hidden trade constraints with the FOC of privately optimal contracts without hidden trade.[40] This highlights two wedges, arising from hidden trade, which work in opposite directions and can be expressed in terms of the precautionary saving motive and the risk premium

[39] In the cashless limit $\beta \to 0$, the IC curve shifts down so that it touches $\hat{x} = a - \rho$ at $\sigma_c/\sigma = 1$ (in general it touches $a - \rho$ at $1 - \beta$). This is of course the non-monetary equilibrium. There may also be another intersection between $\hat{x}^{CE}$ and the IC curve for $\sigma_c/\sigma < 1$, but this violates the No-Ponzi conditions. It's the cashless limit and corresponds to a bubble equilibrium of the non-monetary economy. See section 5 on bubbles and the discussion on the cashless limit in section 2.4.

[40] The HJB equation for the optimal contract without hidden trade (and cashless limit) is $rJ(U) = \min_{c,k} c - k\alpha + J(U)\left(\rho(\rho U - \log c) + \frac{1}{2}(\sigma\rho k/c)^2\right)$, where U is the agent's promised continuation utility and $J(U)$ the cost to the principal. See Di Tella and Sannikov (2016) for details.

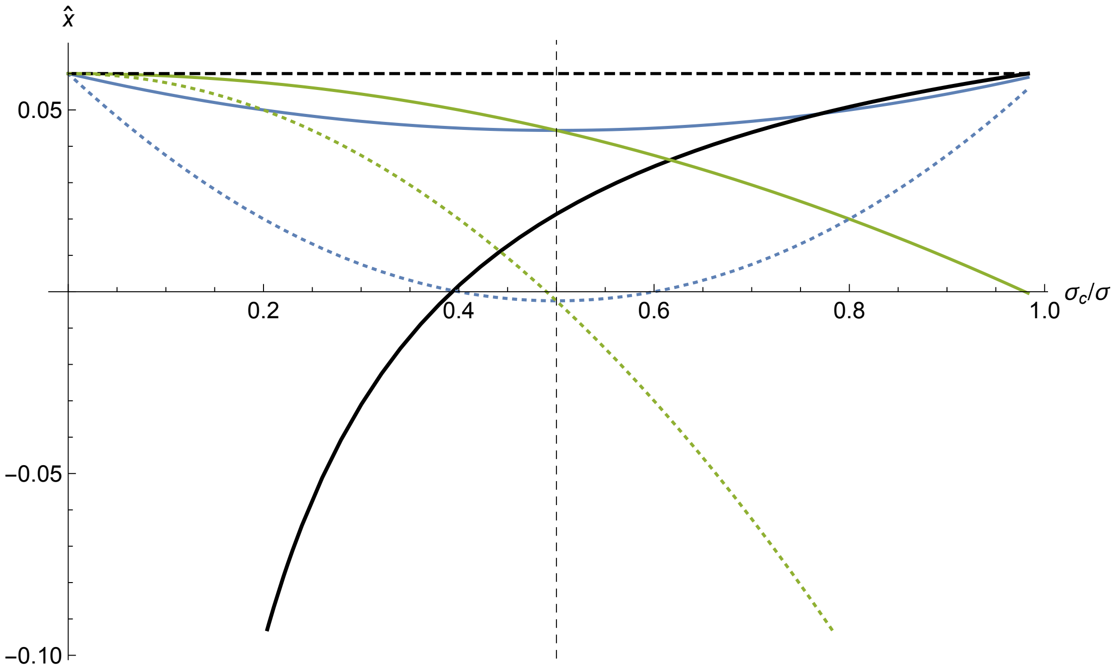

Figure 5: The thick black line is the locus of IC $(\hat{x}, \sigma_c)$. The blue line is the $\hat{x}$ in the CE as a function of σ_c/σ corresponding to (30), the green line for the SP corresponding to (29). Solid is for $\sigma = 0.2$, dotted for $\sigma = 0.5$. *Parameters:* $a = 1/10$, $\rho = 4\%$, $\pi = 2\%$, $\delta = 1\%$, $\beta = 1.7\%$

on capital. The precautionary savings motive generates overinvestment, while an excessive risk premium generates underinvestment.

First, without hidden trade, instead of the Euler equation we have the Inverse Euler equation,

$$\mu_c = r - \rho \qquad \text{Inv. Euler eq. (without hidden trade)} \tag{32}$$

$$\mu_c = r - \rho + \sigma_c^2 - \tau_s \qquad \text{Euler eq. (with hidden trade)} \tag{33}$$

The difference between the Euler and Inverse Euler equations is the precautionary savings motive, σ_c^2. Agents save to self insure, hoping to reduce their exposure to risk σ_c. But their exposure to risk is governed by the IC constraint. Although the IC constraint will ultimately be satisfied, the Euler equation is based on a margin that is not incentive compatible. The correct intertemporal FOC is the Inverse Euler equation that equalizes the cost of delivering utility to the agent across time, taking the IC constraint into account.

The precautionary saving motive generates excessive investment in the competitive equilibrium, relative to the planner's allocation. We can introduce a tax on savings in the environment with hidden trade, $\tau_s = \sigma_c^2$, to transform the Euler equation into the Inverse Euler equation and eliminate the inefficiency.[41]

Hidden trade creates a second inefficiency. With hidden trade agents cannot issue any equity

[41] Put aside for now the question of whether such a tax is compatible with the hidden trade environment. It should be interpreted as a wedge that helps understand the inefficiency created by hidden trade.

and must bear all the risk in capital. The reason is that they equalize the marginal utility of consumption and wealth. If they secretly steal a dollar and consume it, they must lose a full dollar in wealth to align incentives. But without hidden trade, optimal contracts can front-load the agent's consumption to drive the marginal utility of consumption below the marginal utility of wealth. The agent would like to deviate and secretly save, but cannot. Now he can be allowed to issue some equity and share risk, retaining a fraction $\phi < 1$ of the equity. If he secretly steals a dollar and consumes it, he loses only $\phi < 1$ dollars in wealth, but the contract is still incentive compatible because wealth is more valuable than consumption.

Being able to issue equity and share risk makes capital more attractive. The FOC for capital with and without hidden trade is

$$\alpha = a - \delta - r = \sigma_c(\phi\sigma) \qquad \phi \text{ retained equity share (without hidden trade)} \tag{34}$$

$$\alpha = a - \delta - r = \sigma_c\sigma - s_k \qquad \phi = 1 \text{ retained equity share (with hidden trade)} \tag{35}$$

where $\phi = \rho k/c < 1$ is the agent's retained equity share. Ultimately the IC constraint will be satisfied in both cases, but privately optimal contracts with hidden trade generate a risk premium that is too high, and yield underinvestment in the competitive equilibrium relative to the planner's allocation. We can introduce a subsidy to capital $s_k = \sigma_c\sigma - \sigma_c^2$ to transform the FOC for capital with hidden trade into the FOC for capital without hidden trade, so that agents invest as if they could issue some outside equity (when they really cannot).

Looking at the FOCs of the optimal contract without hidden savings helps understand the inefficiencies created by hidden trade, and it's easy to see that if we combine them we obtain the planner's level of investment (29).[42] The two wedges come from hidden trade, but go in opposite directions. The precautionary motive generates over investment, while the excessive risk premium generates underinvestment. Which one dominates depends on the level of idiosyncratic risk σ, as formalized in Proposition 9. Without money, the precautionary motive always dominates, so we have overinvestment for any σ. Money improves risk sharing and weakens both the precautionary motive and the risk premium, but it weakens the precautionary motive more, as explained in section 2.4. As a result, when the value of money is very large, the excessive risk premium dominates and we have underinvestment.

Finally, away from the cashless limit, $\beta > 0$, there is another inefficiency because even the FOCs of private contracts without hidden savings treat real money balances as a scarce resource. However, for a given nominal interest rate i, if all agents increase their consumption the aggregate supply of real money balances also increases, $m_t = \frac{\beta}{1-\beta}\frac{c_t}{i_t}$. Privately optimal contracts don't internalize this.[43]

[42] However, as it turns out, the competitive equilibrium without hidden trade doesn't exist. There is a sequence of global deviations that yield infinite profits for the principal. The setting without hidden trade is only useful to understand the inefficiency in our setting with hidden trade, where the competitive equilibrium is well defined. The planner's solution is well defined in both cases. See Di Tella and Sannikov (2016) for details.

[43] If instead of taking nominal interest rates i as given, we took the path for real money balances m as given in the planner's problem, this inefficiency wouldn't arise.

3.5 Implementing the optimal allocation

We can implement the planner's allocation with a tax on savings and a subsidy to capital, rebated with lump-sum transfers. Alternatively, we can combine both in a tax on capital,[44]

$$\tau_k = \tau_s - s_k = 2\sigma_c^2 - \sigma_c\sigma \tag{36}$$

which is positive or negative depending on which force dominates. Compared to the unregulated competitive equilibrium, total wealth now includes not only money, but also the lump-sum rebates $w_t = k_t + m_t + h_t + \int_t^\infty e^{-rs}\tau_s^k k_s ds$. The only equilibrium condition that changes is the asset pricing equation for capital (6), which becomes

$$\underbrace{a - \tau^k - \delta - r}_{\alpha} = \sigma_c\sigma \tag{37}$$

The competitive equilibrium still has an Euler equation $r = \rho + (\hat{x} - \delta) - \sigma_c^2$, and idiosyncratic risk sharing is given by (26).

We can combine the Euler equation and the FOC for capital to obtain the planner's investment level (29). This allows us to implement the planner's allocation as a competitive equilibrium, provided that the required real interest rate is not too low. Of course, we also need to choose the inflation target to deliver the Friedman rule $i = r + \pi \approx 0$.

A *Balanced Growth Path Equilibrium with tax τ^k and inflation* π is an interest rate r, investment $\hat{x}$, and real money $\hat{m}$ satisfying the Euler equation (5), the asset pricing equation (37), risk sharing equation (26), and money demand equation (9), as well as $i = r + \pi > 0$ and $r > \hat{x} - \delta$.

Proposition 10. *Let $\hat{x}$ be investment in the planner's allocation, with associated idiosyncratic consumption risk σ_c. If $r = \rho + \hat{x} - \delta - \sigma_c^2 > \hat{x} - \delta$, and $r + \pi > 0$, then r, $\hat{x}$, and $\hat{m} = \frac{\beta}{1-\beta}\frac{a-\hat{x}}{r+\pi}$ are a BGP equilibrium with tax τ^k given by* (36) *and inflation π. The optimal inflation target π satisfies $i = r + \pi \approx 0$ (Friedman rule).*

When idiosyncratic risk σ is low, the competitive equilibrium has too much investment and too little risk sharing. The optimal allocation requires a tax on capital $\tau^k > 0$. But when idiosyncratic risk σ is high, the competitive equilibrium has too little investment and too much risk sharing. The optimal allocation then requires a subsidy to capital $\tau^k < 0$.

Notice that $r > \hat{x} - \delta$ also ensures that the privately optimal contract that takes prices r, α, and i as given can be implemented as the consumption-portfolio problem from section 2. We know that $r - (\hat{x} - \delta) = \rho - \sigma_c^2$ falls with σ from Proposition 8, so there is a maximum $\bar{\sigma} = 2\sqrt{\rho}\frac{1}{1-\beta}$ such that the optimal allocation can be implemented as a competitive equilibrium with a tax τ^k for all $\sigma \in [0, \bar{\sigma})$.

[44] This is consistent with the hidden trade in the environment. The planner is taxing or subsidizing all capital, regardless of who holds it.

Proposition 11. *The optimal allocation can be implemented as a BGP with tax τ^k for all $\sigma \in [0,\bar{\sigma})$, with $\bar{\sigma} = \frac{2\sqrt{\rho}}{1-\beta}$. If $\beta < \frac{1}{2}$, there is a $\sigma^* = 2\sqrt{\rho(1-2\beta)} \in (0,\bar{\sigma}]$ such that for $\sigma \leq \sigma^*$ the implementation of the optimal allocation requires $\tau^k \geq 0$, and for $\sigma^* < \sigma < \bar{\sigma}$ it requires a $\tau^k < 0$. If $\beta \geq \frac{1}{2}$, we have $\tau^k \leq 0$ for all $\sigma \in [0,\bar{\sigma})$.*

4 Dynamic model with risk shocks

In this section I incorporate aggregate risk shocks that increase idiosyncratic risk σ_t. I also add TFP shocks that affect the level of effective capital. This dynamic extension allows me to obtain a general formula for the value of liquidity λ that highlights the role of the stochastic discount rate—the asset pricing of liquidity plays a prominent role. It also clarifies the role of idiosyncratic and aggregate risk. What matters about money is that it improves *idiosyncratic* risk sharing; aggregate risk sharing is complete.

4.1 Setting

The "capital quality" shock for an agent is now

$$\Delta^k_{i,t} = \sigma_t k_{i,t} dW_{i,t} + \tilde{\sigma}^{TFP} dZ^{TFP}_t \tag{38}$$

where Z^{TFP} is an aggregate TFP shock. Aggregate TFP risk $\tilde{\sigma}^{TFP}$ is constant, but idiosyncratic risk σ_t follows an autoregressive process

$$d\sigma_t = \phi(\bar{\sigma} - \sigma_t)dt + \sqrt{\sigma_t}\nu dZ^{RS}_t \tag{39}$$

where Z^{RS} is another aggregate shock that raises idiosyncratic risk. For theoretical clarity, assume that Z^{RS} and Z^{TFP} are independent.

Idiosyncratic shocks wash away in the aggregate, so the aggregate capital stock follows

$$dk_t = (gk_t + x_t - \delta k_t)dt + k_t \tilde{\sigma}^{TFP} dZ^{TFP}_t \tag{40}$$

where g is the exogenous growth rate of TFP. Markets are complete for aggregate risk, with endogenous price θ^{TFP}_t and θ^{RS}_t for Z^{TFP} and Z^{RS} respectively. Let Q be the equivalent martingale measure associated with these prices.

As in the baseline model, the government prints money and distributes it lump-sum to agents in order to hit the inflation target. Because we know that the inflation target itself has no real effects (besides real money holdings) we can allow the inflation target π_t to depend on the whole history of aggregate shocks Z^{TFP} and Z^{RS}. For example, it may follow the Friedman rule such that $i_t = r_t + \pi_t \approx 0$. To achieve its inflation target, the government must be ready to introduce or take out money in response to aggregate shocks to accommodate demand for money so that nominal prices grow at rate π_t:[45]

$$\frac{dM_t}{M_t} = \mu_{Mt}dt + \tilde{\sigma}_{Mt}^{TFP} d\tilde{Z}_t^{TFP} + \tilde{\sigma}_{Mt}^{RS} d\tilde{Z}_t^{RS}$$

Let h_t be the present value of future money transfers: $h_t = \mathbb{E}_t^Q \left[\int_t^\infty e^{-\int_t^s r_u du} \frac{dM_s}{p_s} \right]$. Total wealth is $w_t = k_t + m_t + h_t$.

Just as in the baseline model, the value of liquidity plays a central role. But instead of working with $\lambda_t = \frac{m_t + h_t}{k_t + m_t + h_t}$, it is easier to work with $\psi_t = \hat{m}_t + \hat{h}_t$, with law of motion

$$d\psi_t = \mu_{\psi t}dt + \tilde{\sigma}_\psi^{RS} d\tilde{Z}_t^{RS} \tag{41}$$

ψ_t is not exposed to TFP shocks because the economy is scale invariant to effective capital k_t. Total wealth is $w_t = k_t(1 + \psi_t)$, and we can recover $\lambda_t = \frac{\psi_t}{1+\psi_t}$.

With this definition of ψ, the competitive equilibrium is a process for the real interest rate r and price of risk θ^{TFP} and θ^{RS}, investment $\hat{x}$, and real money holdings $\hat{m}$, all contingent on the history of shocks Z^{TFP} and Z^{RS} and satisfying the equilibrium conditions:

$$\begin{aligned}
r_t &= \rho + (\hat{x}_t + g - \delta) + \mu_{\psi t}/(1 + \psi_t) - \sigma_{ct}^2 - (\tilde{\sigma}_{ct}^{TFP})^2 - (\tilde{\sigma}_{ct}^{RS})^2 && \text{euler} && (42)\\
r_t &= a + g - \delta - \sigma_{ct}\sigma_t - \theta_t^{TFP}\tilde{\sigma}^{TFP} && \text{asset pricing} && (43)\\
\sigma_c &= \sigma_t/(1 + \psi_t) && \text{idiosyncratic risk} && (44)\\
\tilde{\sigma}_c^{TFP} &= \theta_t^{TFP} = \tilde{\sigma}^{TFP} && \text{TFP risk} && (45)\\
\tilde{\sigma}_{ct}^{RS} &= \theta_t^{RS} = \tilde{\sigma}_{\psi t}^{RS}/(1 + \psi_t) && \text{RS risk} && (46)\\
\hat{m}_t &= \beta/(1 - \beta) \times (a - \hat{x}_t)/(r_t + \pi_t) && \text{money} && (47)
\end{aligned}$$

as well as $\lim_{t\to\infty} \mathbb{E}^Q \left[e^{-\int_0^t r_u du} w_t \right] = 0$ and $i_t = r_t + \pi_t > 0$.

There is now a precautionary motive for idiosyncratic risk σ_{ct}^2, for aggregate TFP risk $(\tilde{\sigma}_{ct}^{TFP})^2$, and for aggregate risk shocks $(\tilde{\sigma}_{ct}^{RS})^2$. In addition, the growth rate of consumption in the Euler equation $\mu_{ct} = \hat{x}_t + g - \delta + \mu_{\psi t}/(1 + \psi_t)$ comes from writing $c_t = \rho(1-\beta)w_t$ and $w_t = k_t(1+\psi_t)$, and computing the drift of c_t. Likewise, capital pays a risk premium for its idiosyncratic risk $\sigma_{ct}\sigma_t$ and for aggregate TFP risk $\theta_t^{TFP}\tilde{\sigma}^{TFP} = (\tilde{\sigma}^{TFP})^2$ (its return is not correlated with aggregate risk shocks, so there's no risk premium for that). Idiosyncratic risk sharing is still given by $\sigma_{ct} = (1 - \lambda_t)\sigma = \frac{\sigma}{1+\psi_t}$. Money provides a safe store of value that improves idiosyncratic risk sharing. But aggregate shocks cannot be shared. Since the economy is scale invariant to effective capital, TFP shocks don't affect $\hat{x}_t = x_t/k_t$ or $\hat{c}_t = c_t/k_t$, so $\tilde{\sigma}_{ct}^{TFP} = \tilde{\sigma}^{TFP}$. In contrast, risk shocks don't affect the level of effective

[45] Here while the inflation rate π_t can be stochastic, the price level is locally deterministic (no volatility). This simplifies things so that, for example, $i = r + \pi$ holds in equilibrium (otherwise we would need to take a risk premium for nominal bonds into account). But since money is neutral, we could also allow for volatility in the price level without changing the main results. What matters about money is that it does not have idiosyncratic risk; aggregate risk sharing is already complete.

capital, but can affect the value of liquidity and therefore $\hat{x}_t$ and $\hat{c}_t$. So the price of the aggregate risk shock $\theta_t^{RS} = \tilde{\sigma}_{ct}^{RS}$ is endogenous and depends on how the shock affects ψ_t. Equation (46) comes from writing $c_t = \rho(1-\beta)w_t$ and $w_t = k_t(1+\psi_t)$ and computing the exposure of consumption to aggregate risk shocks.

If we know the behavior of ψ_t we can then obtain every other equilibrium object from (42)-(47). In contrast to the baseline model, the value of liquidity is not a constant. It will be characterized by an ODE. Reasoning as before, the value of money is equal to the present value of expenditures on liquidity services.

Proposition 12. *The equilibrium value of money satisfies*

$$m_t + h_t = \mathbb{E}_t^Q \left[\int_t^\infty e^{-\int_t^s r_u du} m_s i_s ds \right] \tag{48}$$

From this it is easy to obtain a BSDE for ψ_t

$$\hat{m}_t i_t + \mu_{\psi t} + \psi_t(\hat{x}_t + g - \delta - r_t) = \theta_t^{TFP} \tilde{\sigma}^{TFP} \psi_t + \theta_t^{RS} \tilde{\sigma}_{\psi t}^{RS} \tag{49}$$

and transversality condition $\lim_{T\to\infty} \mathbb{E}^Q \left[e^{-\int_0^T r_u du} k_T \psi_T \right] = 0$.

It is worth noting that because of the log preferences, dynamics only matter through the value of liquidity ψ. Without money we have $\psi_t = 0$, and the competitive equilibrium does not depend on the stochastic process for σ_t, as can be seen from inspecting equations (42)-(47) (only on the current σ_t). For any given value of σ_t the competitive equilibrium without money is the same as in the static economy with a constant $\sigma = \sigma_t$. With money, the only reason this is not the case is because ψ_t is forward-looking, as shown in equation (48). As we will see below, the planner's optimal allocation for a given value of σ_t also coincides with the static case with a constant $\sigma = \sigma_t$.

We can subtract equation (42) from (43) and re-arrange to obtain an expression for investment $\hat{x}_t$, analogous to (15):

$$\hat{x} = a - \rho + \sigma_{ct}^2 - \sigma_{ct}\sigma + \left(\frac{\tilde{\sigma}_{\psi t}^{RS}}{1+\psi_t} \right)^2 - \frac{\mu_{\psi t}}{1+\psi_t} \tag{50}$$

Idiosyncratic and aggregate risk play very different roles. Money provides a safe store of value that improves idiosyncratic risk sharing. As before, it dampens the idiosyncratic precautionary motive σ_{ct}^2 relative to the idiosyncratic risk premium $\sigma_{ct}\sigma$ and depresses investment. Aggregate risk, on the other hand, simply cannot be shared. As a result, aggregate TFP risk $\tilde{\sigma}^{TFP}$ reduces the equilibrium real interest rate r, but does not affect investment $\hat{x}$, just as in the non-monetary economy. The risk premium $\tilde{\sigma}_{ct}^{TFP} \tilde{\sigma}^{TFP}$and precautionary motive $(\tilde{\sigma}_{ct}^{TFP})^2$ produced by aggregate TFP risk exactly cancel each other out even when there is money. Likewise, the exogenous growth rate of TFP g affects only the equilibrium real interest rate r, but has no effect on the value of liquidity ψ or investment $\hat{x}$.[46]

[46] This is a property of preferences with EIS of one. In general $\tilde{\sigma}^{TFP}$ and g could affect investment $\hat{x}$ even without money, and can affect the value of liquidity ψ because $r + \theta^{TFP}\tilde{\sigma}^{TFP} - g$ is not invariant to changes in g and $\tilde{\sigma}^{TFP}$.

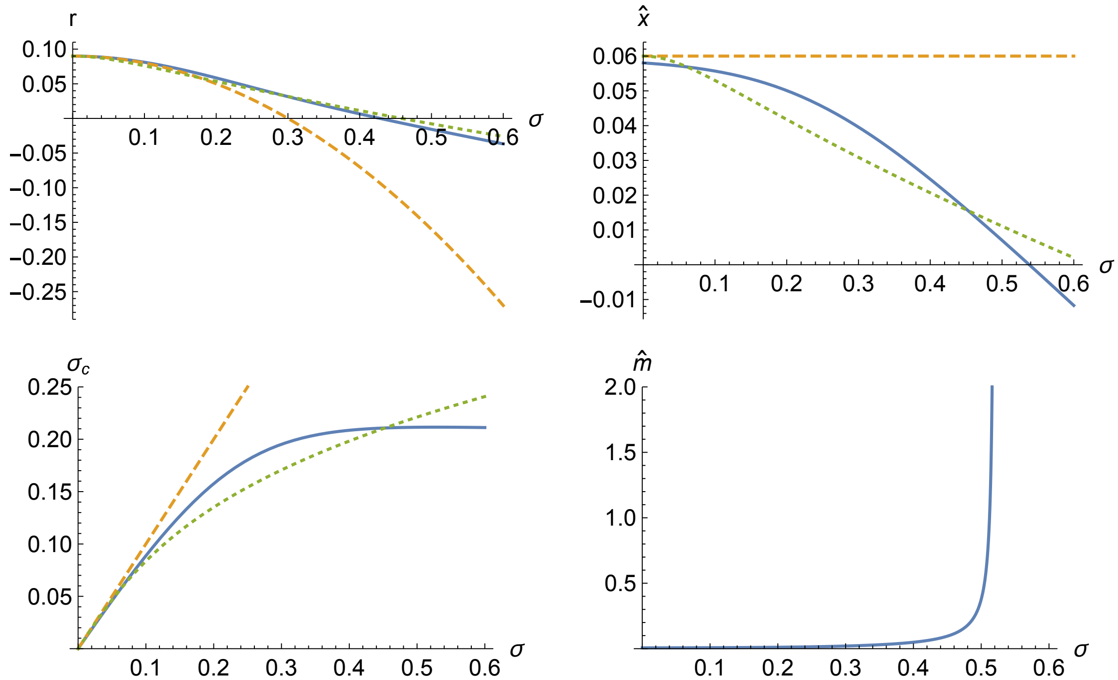

Figure 6: Interest rate r, investment $\hat{x}$, idiosyncratic risk σ_c, and real money balances $\hat{m}$ in the dynamic model for the non-monetary economy (dashed orange), monetary competitive equilibrium (solid blue), and the social planner's allocation (dotted green). *Parameters:* $a = 1/10$, $\rho = 4\%$, $\delta = 1\%$, $\beta = 1.7\%$, $\bar{\sigma} = 0.3$, $\phi = 0.02$, $\nu = 0.01$, $\tilde{\sigma}^{TFP} = 0$. Real money balances are computed with $\pi = 2\%$.

The stochastic behavior of σ_t matters through the equilibrium behavior of the value of liquidity ψ_t, which is forward-looking. Even conditional on the current value of ψ_t, if ψ_{t+s} is expected to be high in the future, this means the idiosyncratic precautionary motive will be weaker than the idiosyncratic risk premium and therefore consumption will be higher relative to capital; that is, $\hat{c}_{t+s}$ will be higher. As a result, for a given interest rate r_t, pinned down by the asset pricing equation (43), agents want less investment $\hat{x}_t$ to achieve their desired intertemporal consumption smoothing. This is why $\mu_{\psi t}$ appears in (50). Likewise, aggregate risk shocks matter because they induce aggregate volatility in agents' consumption. Since capital is not exposed to risk shocks, it is an attractive hedge. This is why $\tilde{\sigma}^{RS}_{\psi t}$ appears in (50).

Recursive Equilibrium. We look for a recursive equilibrium with σ_t as the state variable, so $\psi_t = \psi(\sigma_t)$ can be characterized as the solution to an ODE derived from (49). Use Ito's lemma to compute the drift and volatility of ψ, the Euler equation (42) to eliminate r and $\hat{x}$ terms, and $\hat{m}_t i_t = \rho\beta w_t/k_t = \rho\beta(1+\psi_t)$.

Proposition 13. *The equilibrium value of $\psi(\sigma)$ solves the ODE*

$$\rho\beta(1+\psi) + \frac{\psi'\phi(\bar{\sigma}-\sigma) + \frac{1}{2}\psi''\sigma\nu^2}{1+\psi} = \psi(\rho - (\frac{\sigma}{1+\psi})^2) + \left(\frac{\psi'\sqrt{\sigma}\nu}{1+\psi}\right)^2 \tag{51}$$

With a solution $\psi(\sigma)$ to (51) we can obtain all the other equilibrium objects using (42)-(47). If we also satisfy the transversality conditions, we have a competitive equilibrium.

Figure 6 shows the competitive equilibrium in the dynamic model, with and without money. It has essentially the same properties as in the BGP economy with constant idiosyncratic risk σ. Money improves idiosyncratic risk sharing but depresses investment. After a risk shock increases σ_t the value of liquidity goes up, so the real interest rate falls less than it would without money, and investment falls instead. The effects of the shock are as persistent as the shock itself, and inflation is always on target π_t. The target itself doesn't affect any real variable other than real money holdings m_t through the interest rate, and this is valid for time-varying inflation targets. As a result, the optimal inflation target satisfies the Friedman rule $i_t = r_t + \pi_t \approx 0$, so π_t must go up during periods of high idiosyncratic risk σ_t (because the real interest rate r_t is then low).[47]

4.2 Planner's problem

As in the static case, the incomplete idiosyncratic risk sharing can be microfounded by a moral hazard problem with hidden trade (see Online Appendix for details). The incentive compatibility constraints are

$$\sigma_{ct} = \rho(1-\beta)c_t^{-1}k_t\sigma \qquad \text{"skin in the game"} \qquad (52)$$

$$\mu_{ct} = r_t - \rho + \sigma_{ct}^2 + (\tilde{\sigma}_{ct}^{TFP})^2 + (\tilde{\sigma}_{ct}^{RS})^2 \qquad \text{Euler equation} \qquad (53)$$

$$\alpha_t = \sigma_{ct}\sigma_t \qquad \text{demand for capital} \qquad (54)$$

$$\tilde{\sigma}_{ct}^{TFP} = \theta_t^{TFP} \qquad \text{TFP shocks} \qquad (55)$$

$$\tilde{\sigma}_{ct}^{RS} = \theta_t^{RS} \qquad \text{risk shocks} \qquad (56)$$

$$m_t/c_t = \beta/(1-\beta)i_t^{-1} \qquad \text{demand for money} \qquad (57)$$

As before, the planner does not take prices r, α, θ^{TFP}, θ^{RS}, and i as given, so the only true IC constraint is that he must treat all agents the same (up to scale) and must expose them to their own idiosyncratic risk to provide incentives. This yields a tradeoff between investment $\hat{x}_t$ and idiosyncratic risk sharing σ_{ct}/σ_t

$$\sigma_{ct}/\sigma_t = \frac{\rho(1-\beta)}{a - \hat{x}_t} \qquad (58)$$

which is the same as in the static version of the model given by (31).

Analogous to the one in the baseline setting, we use the fact that all agents must be treated the same up to scale to write the objective function for the planner as

$$\mathbb{E}\left[\int_0^\infty e^{-\rho t}\left(\log k_0 + \log(a - x_t) + \frac{x_t - \delta}{\rho} - \frac{1}{2}\frac{(\tilde{\sigma}^{TFP})^2}{\rho} + \beta\log\left(\frac{\beta}{1-\beta}i_t^{-1}\right)\right)dt\right]$$

[47] It's worth noting that the central bank might be unable to deliver a fixed inflation target $\pi_t = \bar{\pi}$, because for high enough σ_t the real interest rate becomes very negative. When this happens the central bank is forced to abandon its inflation target. The competitive equilibrium with money only exists for inflation targets π_t that take this into account (such as $\pi_t = \max\{\tilde{\pi}_t, -r_t\}$ for an arbitrary $\tilde{\pi}_t$), but the inflation target doesn't matter otherwise for the real side of the economy.

As in the baseline, for an arbitrary path for nominal interest rates i, we solve for the optimal allocation of capital and consumption subject to the IC constraint (58). The solution trades off consumption, growth, and risk sharing and is independent of the path for i. Given this optimal allocation, the optimal monetary policy is to follow the Friedman rule, $i \approx 0$. The FOC for $\hat{x}$ yields

$$\hat{x}_t = a - \rho - \sigma_{ct}^2 \tag{59}$$

which is the same as in the stationary baseline model. The source of inefficiency is the same as in the stationary model, but it now has a dynamic dimension. As before, the planner and the competitive equilibrium share the same locus of incentive-compatible combinations of investment $\hat{x}$ and idiosyncratic risk sharing σ_c given by (58), but they choose different $(\hat{x}, \sigma_c)$ combinations. However, while the planner only cares about the current value of idiosyncratic risk σ_t, the competitive equilibrium picks a $(\hat{x}_t, \sigma_{ct})$ that depends on the future behavior of idiosyncratic risk σ_t through the value of liquidity today ψ_t.

Figure 6 also shows the planner's optimal allocation. An important difference with the static case is that it is possible for investment to be too low in the competitive equilibrium for very low σ_t, too high for intermediate σ_t, and again too low for high σ_t (and the same goes for σ_c through (58)). The reason for this is that while the planner's allocation coincides with the static case with a constant $\sigma = \sigma_t$, the competitive equilibrium depends on the stochastic behavior of σ_t through the value of liquidity ψ_t. If σ_t today is very low but is expected to go up to $\bar{\sigma}$ very fast, the value of ψ_t will be closer to the value in the static case with $\sigma = \bar{\sigma}$ rather than $\sigma = \sigma_t$. So the competitive equilibrium chooses a $(\hat{x}_t, \sigma_{ct})$ on the incentive-compatibility constraint given by (58) that is closer to what it would choose in the static case with $\sigma = \bar{\sigma}$. This is captured in equation (50) through the role of $\mu_{\psi t}$ and $\sigma_{\psi t}$.[48]

We can still implement the planner's optimal allocation with a tax on capital τ^k that internalizes the externality produced by hidden trade. Introducing the tax only changes the asset pricing equation

$$r_t = a + g - \delta - \tau_t^k - \sigma_{ct}\sigma_t - \theta_t^{TFP}\tilde{\sigma}^{TFP}$$

and therefore the equilibrium investment

$$\hat{x}_t = a - \rho - \tau_t^k + \sigma_{ct}^2 - \sigma_{ct}\sigma_t - \mu_{\hat{c}t} + (\tilde{\sigma}_{ct}^{RS})^2$$

Comparing this expression to (59) we find the wedge

$$\tau_t^k = 2\sigma_{ct}^2 - \sigma_{ct}\sigma_t - \mu_{\hat{c}t} + (\tilde{\sigma}_{ct}^{RS})^2 \tag{60}$$

Proposition 14. *Let $\mathcal{P}$ be an optimal allocation with processes for investment $\hat{x}$, aggregate consumption $\hat{c}$, and idiosyncratic consumption risk σ_c. Then $r_t = \rho + \mu_{ct} - \sigma_{ct}^2 - (\tilde{\sigma}_{ct}^{TFP})^2 - (\tilde{\sigma}_{ct}^{RS})^2$, $\theta_t^{TFP} = \tilde{\sigma}_c^{TFP}$, $\theta_t^{RS} = \tilde{\sigma}_c^{RS}$, $\hat{x}$ and $\hat{m}_t = \frac{\beta}{1-\beta}\frac{a-\hat{x}_t}{r_t+\pi_t}$ is a competitive equilibrium with τ_t^k given by*

[48] In this solution, $\sigma_\psi \to 0$ as $\sigma \to 0$, so it's all really about the drift μ_ψ, which does not vanish as $\sigma \to 0$.

(60) *and inflation target* π_t*, provided that* $\lim_{t\to\infty}\mathbb{E}^Q\left[e^{-\int_0^t r_u du}k_t\hat{c}_t\right]=0$ *and* $i_t=r_t+\pi_t>0$*. The optimal inflation target satisfies* $i_t=r_t+\pi_t\approx 0$ *(Friedman rule).*

5 Discussion

5.1 Liquidity premium and bubbles

In the model the liquidity premium comes from the fact that liquid assets appear in the utility function, so agents are willing to hold them even if their yield is below the interest rate. This is the simplest and most transparent way of introducing assets with a liquidity premium, and it is meant to capture that money and other assets are useful for transactions. I also solve the model with a cash-in-advance constraint and a more flexible CES specification for MIU, and obtain the same results except that the expenditure share on liquidity β now depends on the nominal interest rate. We could also use a more microfounded model of monetary exchange, as in Lagos and Wright (2005). The liquidity premium can also reflect the use of some securities, such as Treasuries, in financial transactions, or their role in financial regulation (e.g. capital requirements). In fact, at least part of the liquidity premium on Treasuries is likely to reflect this.[49] As long as money and other safe assets have a liquidity premium—that is, their yield is below the risk-free rate—the results will go through.

There is also a large literature modeling money as a bubble in the context of OLG or incomplete risk sharing models.[50] The closest paper is Brunnermeier and Sannikov (2016b) who develop a version of the Bewley (1980) model that is tractable and yields closed-form solutions. In contrast, here bubbles are explicitly ruled out.[51] The liquidity view in this paper has differences and similarities with the bubble view. Many of the results here are likely to go through in a setting with bubbles. Intuitively, a safe bubble can provide a safe store of value that improves idiosyncratic risk sharing and depresses investment, just like money does in this paper. But there are also important differences.

To understand the link with bubbles, we can ignore the government bonds and deposits, $b_t = d_t = 0$, and assume money is not printed, $dM_t = 0$, and therefore in a BGP inflation is simply

[49] Krishnamurthy and Vissing-Jorgensen (2012) show that US Treasuries have a convenience yield over AAA corporate debt. They further decompose the convenience yield into a liquidity premium and a safety premium. Here I don't draw this distinction, and call the whole spread a liquidity premium. For the purpose of the mechanism in the paper the exact origin of this premium is not essential.

[50] See Samuelson (1958), Bewley (1980), Aiyagari (1994), Diamond (1965), Tirole (1985), Asriyan et al. (2016), Santos and Woodford (1997).

[51] In addition, they introduce money proportionally to wealth, so high inflation acts as a subsidy to saving. They find that the optimal inflation rate rises with risk. Here instead money is distributed lump-sum, so it is superneutral—changes in the inflation rate don't have real effects. The model here converges to theirs in the cashless limit $\beta\to 0$ in the special case where money supply is constant, $\mu = 0$.

minus the growth rate of the economy. Write equation (18) without the No-Ponzi condition

$$w_t = k_t + \underbrace{\int_t^{\infty} e^{-\int_t^s r_u du} m_t i_t dt}_{\text{liquidity services}} + \underbrace{\lim_{T\to\infty} e^{-\int_0^T r_s ds} m_T}_{\text{bubble}}$$

In this paper, the No-Ponzi condition eliminates the last term, $\lim_{T\to\infty} e^{-\int_0^T r_s ds} m_T = 0$. Money is not a bubble—it derives its value from its liquidity premium, which arises because money provides liquidity services. In models with bubbles, instead, money doesn't have a liquidity premium, so the nominal interest rate must be zero $i_t = 0$. Money is an asset that doesn't pay any dividend, but still yields the arbitrage-free market return; r_t if there is no aggregate risk. It has positive value only because the last term doesn't vanish. In a BGP, this requires $r =$ growth rate. But in both the liquidity and the bubble views, money provides a store of value that improves risk sharing, $\sigma_c = \frac{k_t}{w_t}\sigma < \sigma$.

Modeling money as an asset with a liquidity premium has several advantages. First, money does have a liquidity premium. The bubble view cannot explain why people hold money when they can hold safe nominal bonds that pay interest. If the bubble is really money, then the equilibrium nominal interest rate must be zero and the real interest rate must be equal to the growth rate of the economy. The liquidity view instead can provide a more flexible account of inflation and interest rates. In addition, the bubble only exists for high enough levels of idiosyncratic risk, while here money always has equilibrium real effects which gradually grow as interest rates fall.

Alternatively, the bubble may not really be money. It could be housing, the stock market, government debt, social security, or even tulips. This can potentially be very interesting, but it is difficult to determine if asset values really have a bubble component. In contrast, it's relatively straightforward to establish that some assets have a liquidity premium. Linking the value of money to liquidity premiums also allows us to understand its behavior in response to shocks and policy interventions, since it is more readily grounded in fundamentals.

Finally, it's worth pointing out that bubbles may have idiosyncratic risk. To the extent that agents cannot diversify this idiosyncratic risk, the bubble will not perform the same role as money, which is safe. For example, suppose there is a housing bubble so that house prices are 10% above their fundamental value, but each agent must buy one house whose value has idiosyncratic risk.

There is also a link between the liquidity view and the bubble view in the cashless limit. As explained in section 2.4, the real effects of money in this paper do not hinge on large expenditures on liquidity services (large β). This can be formalized by noting that the monetary economy does not converge to the non-monetary one as $\beta \to 0$. The reason is that for high enough σ, the real interest rate in the non-monetary economy can be below the growth rate of the economy. This is not a problem in the non-monetary economy, since the only asset is risky capital. But in an economy with safe money, the value of liquidity will become very large as the real interest rate approaches the growth rate of the economy. The result is that the presence of money has real effects even for vanishingly small β. The cashless limit of the monetary economy therefore coincides with a bubbly equilibrium of the non-monetary economy, since it only violates the No-Ponzi condition when $\beta = 0$. For $\beta > 0$ it doesn't violate the No-Ponzi conditions, but gets arbitrarily close.

5.2 Sticky prices and the zero lower bound on interest rates

An alternative view of the role of money focuses on the zero lower bound on nominal interest rates in the context of New Keynesian models with sticky prices. In these models, the natural interest rate r_t^n is the equilibrium real interest rate under flexible prices. The central bank could achieve its inflation target $\bar{\pi}$ and zero output gap by setting the nominal interest rate $i_t = r_t^n + \bar{\pi}$, thus reproducing the flexible-price allocation. This is, in fact, the optimal monetary policy in this environment. But if the natural interest rate is very negative, this could require a negative nominal interest rate. To fix ideas, let the inflation target be zero, $\bar{\pi} = 0$. If the natural rate r_t^n is negative, the central bank is forced to miss its inflation target, allow the economy to operate with an output gap, or both.[52]

Under both views, money prevents the real interest rate from falling to stabilize investment and creates a slump. In the New Keynesian setting because of sticky prices and the ZLB; in this paper because it provides a store of value that improves risk sharing, even though prices are flexible. Both approaches are potentially complementary. In the short run prices may well be sticky, but I abstract from these short run frictions. The model in this paper can be regarded as the flexible-price version of a richer model with nominal rigidities, and the real interest rate here as the natural rate in that model.

This has important consequences for the role of the zero lower bound. First, while the presence of money does create a ZLB, it also raises the natural interest rate. It is perfectly possible for the natural interest rate to be negative without money but positive with money, even if prices are flexible. Under these circumstances the zero lower bound will not be a problem—the central bank will be able to hit its inflation target and a zero output gap; that is, to reproduce the flexible-price allocation. Furthermore, since money is superneutral, a higher inflation target will always "fix" the zero lower bound problem, but it will not affect the real side of the economy. In fact, under the optimal monetary policy, with $i \approx 0$, the zero lower bound is never an issue.

However, the flexible-price equilibrium is only optimal if a tax/subsidy on investment is used to correct the externality produced by hidden trade. If it is not, or if it cannot be adjusted fast enough, reproducing the flexible-price allocation is not optimal. Investment is too depressed when idiosyncratic risk is high. The central bank may then lower the interest rate below the natural level in order to stimulate investment. I conjecture that this is in fact the optimal policy in this environment (with nominal rigidities) but this is beyond the scope of this paper.

[52] See Werning (2011) for a characterization of the optimal monetary policy in a New Keynesian model where the zero lower bound is binding.

5.3 The role of the risk premium

The model in this paper is driven by exogenous shocks to idiosyncratic risk σ_t for the sake of concreteness. But the mechanism actually hinges on the idiosyncratic risk premium and precautionary motives, which might go up for other reasons. For example, suppose we generalize preferences to EZ with intertemporal elasticity 1 but risk aversion $\gamma > 1$ (see Appendix for details). The only thing that changes in the equations is that the risk premium is now $\alpha = \gamma\sigma_c \times \sigma$, and the precautionary motive is $\gamma\sigma_c^2$ (in other words, the "price" of idiosyncratic risk is $\gamma\sigma_c$). We obtain a system of equations analogous to (14), (15), and (8) (with $\gamma = 1$ we recover the original equations):

$$r = a - \delta - \gamma\sigma^2(1-\lambda) \tag{61}$$

$$\hat{x} = a - \rho - \rho\frac{\lambda - \beta}{1-\lambda} \tag{62}$$

$$\lambda = \frac{\rho\beta}{\rho - \gamma\sigma^2(1-\lambda)^2} \tag{63}$$

Clearly, $\gamma\sigma^2$ is the relevant parameter. An increase in risk aversion γ has the same effect as an increase in idiosyncratic risk σ^2. An increase in risk aversion could represent changes in wealth distribution between risk averse and risk tolerant agents after a negative shock (see Longstaff and Wang (2012) or Gârleanu and Panageas (2015)) or the health of the balance sheets of specialists who invest in risky assets such as entrepreneurs or financial intermediaries (see He and Krishnamurthy (2013) and He et al. (2015))). It could also represent habits (see Campbell and Cochrane (1999)) or an increase in ambiguity aversion after a shock that upends agents' beliefs about how the economy works (see Barillas et al. (2009)).

5.4 Measuring the value of liquidity

The value of liquidity λ plays a central role and captures the equilibrium real effects of money on the economy. The model produces a simple formula for λ in terms of fundamentals, and the cashless limit in Proposition 6 ensures that the real effects of money survive even when $\beta \to 0$. But we might also like to measure λ directly. Unfortunately, since λ represents the present value of liquidity, it cannot be directly measured. We must specify a model of the liquidity premium and a model of asset pricing. The covariance of the liquidity premium with the pricing kernel is especially important because it tells us what is the appropriate discount rate. This paper provides such a model, of course, but it is too stylized to get the asset pricing right (e.g. the equity premium puzzle). A quantitatively serious asset pricing model of the value of liquidity is beyond the scope of this paper, but is a natural next step to evaluate this theory.

Alternatively, we can specify a model of government surpluses and use the budget equation of the government,

$$m_t + b_t = \mathbb{E}_t^Q\left[\int_t^\infty e^{-\int_t^s r_u du} d\tau_s\right] + \mathbb{E}_t^Q\left[\int_t^\infty e^{-\int_t^s r_u du}\left(m_s i_s + b_s(i_s - i_s^b)\right) ds\right]$$

to obtain a measure of at least the part of the value of liquidity corresponding to government liabilities (currency, reserves, and government bonds—this leaves out private liquidity). What this budget constraint says is that the current value of government liabilities $m_t + b_t$ must be financed by the present value of government surpluses (the first term), plus the present value of the liquidity premium (the second term, and the object we are interested in). Since $m_t + b_t$ can be observed, if we specify a model of government surpluses and an asset pricing model, we can back out the present value of the liquidity premium. Notice that if we expect large deficits for a long time, we will back out a very large value of liquidity. In other words, if the government has large liabilities and plans to run a deficit, then through the lens of the government's intertemporal budget constraint, it must mean that the value of liquidity is very large.

5.5 The role of shocks to capital

The combination of idiosyncratic shocks to capital quality and homothetic preferences yield a tractable setting that can be solved in closed form. Risky capital is also essential to obtain a time-varying risk premium, which is central to the mechanism under consideration. However, it's worth asking what the role of money would be if we had uninsurable labor income risk instead. The bottom line is that the role of money would be essentially the same, providing a safe store of value that improves risk sharing and keeps interest rates high and investment low, relative to the non-monetary economy. But since there is no risk premium, we wouldn't be able to think of a shock to the risk premium that reduces demand for capital.

To fix ideas, consider a standard Aiyagari economy with uninsurable labor income. Output is produced with capital and labor, $y_t = f(k_t, l_t) = k_t^\alpha l_t^{1-\alpha}$, which can be used for consumption or investment:

$$dk_t = (f(k_t, l_t) - c_t - \delta k_t)dt$$

There's a constant supply of money M. Each agent's labor endowment $e_{it} \in [\underline{e}, \bar{e}]$ follows a well behaved, independent stochastic process, but aggregate labor supply is constant, $\int e_{it} di = l_t = 1$. Agents have log preferences over consumption and money as in the baseline setting. They have access to risk-free bonds with equilibrium interest rate $r_t = f'_k(k_t, 1) - \delta$, and receive an equilibrium wage $w_t = f'_l(k_t, 1)$. Their budget constraint is

$$dn_{it} = (n_{it} r_t + w_t e_{it} - c_{it} - m_{it} i_t)dt$$

and natural borrowing limit $n_{it} \geq \underline{n}_t \equiv -\int_t^\infty e^{-\int_t^s r_u du} \underline{e} w_s ds$.

Focus on the steady state. We can write $k_{ss}(r)$ and $w_{ss}(r)$ using the equilibrium conditions for r and w, and $m_{ss}(r)$ for the demand for real money balances. Market clearing requires that agents hold all the capital and money in the economy.,

$$\int n g(n; r) dn = k_{ss}(r) + m_{ss}(r)$$

where g is the steady state distribution of wealth, conditional on the real interest rate r.

We want to compare the monetary economy with $\beta > 0$ to a non-monetary economy with $\beta = 0$.

Proposition 15. *The monetary economy has a higher steady state real interest rate and a lower steady state capital stock, compared to the non-monetary economy. That is,* $r_{ss}(\beta) > r_{ss}(0)$ *and* $k_{ss}(\beta) < k_{ss}(0)$ *for* $\beta > 0$.

In this economy there is no risk premium because capital is safe. Uninsurable labor income creates a precautionary saving motive that drives the steady state real interest rate down and the capital stock up relative to the economy without risk. Money improves risk sharing and weakens the precautionary saving motive, so the interest rate goes up and capital down relative to the non-monetary economy. Without risk money has no effects: $r = \rho$ and $f_k'(k_{ss}, 1) = \rho + \delta$ regardless of β.

Since there is no risk premium, we cannot study the effects of an increase in the risk premium that depresses demand for capital at given prices. In fact, without money, since capital is perfectly safe, an increase in labor income risk would increase the precautionary savings motive and steady state capital.

6 Conclusions

In this paper I provide a flexible-price theory of the role of money in an economy with incomplete idiosyncratic risk sharing. After an increase in the risk premium reduces demand for capital, money prevents interest rates from falling and depresses investment, relative to the non-monetary economy. The main insight is that because of its liquidity premium, money provides a safe store of value that improves idiosyncratic risk sharing. This drives a wedge between the risk premium on capital and agents' precautionary saving motive, which keeps interest rates high and depresses investment. While the value of money, and other safe assets with a liquidity premium, is pretty small during normal times, during periods of low real interest rates its value can become very large. In fact, the real effects of money survive even in the cashless limit.

The competitive equilibrium is inefficient. During booms with low risk investment is too high and money doesn't provide enough risk sharing; during downturns with high risk investment is too depressed and money provides too much risk sharing. The inefficiency comes from the inability of private contracts to prevent agents from engaging in hidden trade. While the presence of money has large real effects, monetary policy is ineffective. Money is superneutral, Ricardian equivalence holds, and the zero lower bound on nominal rates is not binding. The optimal allocation can be implemented with a tax or subsidy on capital and the Friedman rule. When investment is too high, tax it; when it's too low, subsidize it.

References

Aiyagari, S. R. (1994). Uninsured idiosyncratic risk and aggregate saving. *The Quarterly Journal of Economics*, 659–684.

Aiyagari, S. R. and E. R. McGrattan (1998). The optimum quantity of debt. *Journal of Monetary Economics 42*(3), 447–469.

Aiyagari, S. R. and N. Wallace (1991). Existence of steady states with positive consumption in the kiyotaki-wright model. *The Review of Economic Studies 58*(5), 901–916.

Angeletos, M. (2006, November). Uninsured idiosyncratic investment risk and aggregate saving. forthcoming in Review of Economic Dynamics.

Aruoba, S. B., C. J. Waller, and R. Wright (2011). Money and capital. *Journal of Monetary Economics 58*(2), 98–116.

Asriyan, V., L. Fornaro, A. Martin, and J. Ventura (2016, September). Monetary Policy for a Bubbly World. NBER Working Papers 22639, National Bureau of Economic Research, Inc.

Bansal, R., D. Kiku, I. Shaliastovich, and A. Yaron (2014). Volatility, the macroeconomy, and asset prices. *Journal of Finance*.

Bansal, R. and A. Yaron (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *The Journal of Finance 59*(4), pp. 1481–1509.

Barillas, F., L. P. Hansen, and T. J. Sargent (2009). Doubts or variability? *journal of economic theory 144*(6), 2388–2418.

Bewley, T. (1980). The optimum quantity of money. *Models of monetary economies*, 169–210.

Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica 77*(3), pp. 623–685.

Bloom, N., M. Floetotto, N. Jaimovich, I. Saporta-Eksten, and S. Terry (2012). Really uncertain business cycles. NBER working paper 18245.

Brunnermeier, M. K. and Y. Sannikov (2016a). The i theory of money. Technical report, National Bureau of Economic Research.

Brunnermeier, M. K. and Y. Sannikov (2016b, May). On the optimal inflation rate. *American Economic Review 106*(5), 484–89.

Buera, F. and J. P. Nicolini (2014). Liquidity traps and monetary policy: Managing a credit crunch.

Caballero, R. J. and A. Simsek (2017). A risk-centric model of demand recessions and macroprudential policy. Technical report, National Bureau of Economic Research.

Campbell, J., S. Giglio, C. Polk, and R. Turley (2012, March). An intertemporal capm with stochastic volatility. working paper.

Campbell, J., M. Lettau, B. G., and Y. Xu (2001). Have individual stocks become more volatile? an empirical exploration of idiosyncratic risk. *The Journal of Finance 56*(1), pp. 1–43.

Campbell, J. Y. and J. H. Cochrane (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of political Economy 107*(2), 205–251.

Christiano, L., R. Motto, and M. Rostagno (2014). Risk shocks. *American Economic Review 104*(1).

Cochrane, J. H. (2011). Presidential address: Discount rates. *The Journal of finance 66*(4), 1047–1108.

Cole, H. L. and N. R. Kocherlakota (2001). Efficient allocations with hidden income and hidden storage. *The Review of Economic Studies 68*(3), 523–542.

Di Tella, S. (2016). Optimal regulation of financial intermediaries. Working paper - Stanford GSB.

Di Tella, S. (2017, December). Uncertainty shocks and balance sheet recessions. *Journal of Political Economy 125*(6), 2038–2081.

Di Tella, S. and Y. Sannikov (2016). Optimal asset management contracts with hidden savings. Working paper, Stanford GSB and Princeton University.

Diamond, P. A. (1965). National debt in a neoclassical growth model. *The American Economic Review 55*(5), 1126–1150.

Dotsey, M. and P. D. Sarte (2000). Inflation uncertainty and growth in a cash-in-advance economy. *Journal of Monetary Economics 45*(3), 631–655.

Eggertsson, G. B. et al. (2003). Zero bound on interest rates and optimal monetary policy. *Brookings papers on economic activity 2003*(1), 139–233.

Eggertsson, G. B. and P. Krugman (2012). Debt, deleveraging, and the liquidity trap: A fisher-minsky-koo approach. *The Quarterly Journal of Economics 127*(3), 1469–1513.

Eggertsson, G. B. and M. Woodford (2004). Policy options in a liquidity trap. *The American Economic Review 94*(2), 76–79.

Farhi, E., M. Golosov, and A. Tsyvinski (2009). A theory liquidity and regulation of financial intermediation. *The Review of Economic Studies 76*, 973–992.

Gârleanu, N. and S. Panageas (2015). Young, old, conservative, and bold: The implications of heterogeneity and finite lives for asset pricing. *Journal of Political Economy 123*(3), 670–685.

He, Z., B. T. Kelly, and A. Manela (2015). Intermediary asset pricing: New evidence from many asset classes. *Available at SSRN 2662182*.

He, Z. and A. Krishnamurthy (2013). Intermediary asset pricing. *The American Economic Review 103*(2), 732–770.

Karatzas, I. and S. Shreve (2012). *Brownian motion and stochastic calculus*, Volume 113. Springer Science & Business Media.

Kehoe, T. and D. Levine (1993). Debt-constrained asset markets. *The Review of Economic Studies 60*, 865–888.

Kiyotaki, N. and R. Wright (1993). A search-theoretic approach to monetary economics. *The American Economic Review*, 63–77.

Krishnamurthy, A. and A. Vissing-Jorgensen (2012). The aggregate demand for treasury debt. *Journal of Political Economy 120*(2), 233–267.

Krugman, P. R., K. M. Dominquez, and K. Rogoff (1998). It's baaack: Japan's slump and the return of the liquidity trap. *Brookings Papers on Economic Activity 1998*(2), 137–205.

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

Light, B. (2018, Dec). Uniqueness of equilibrium in a bewley–aiyagari model. *Economic Theory*.

Longstaff, F. A. and J. Wang (2012). Asset pricing and the credit market. *The Review of Financial Studies 25*(11), 3169–3215.

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

Santos, M. S. and M. Woodford (1997). Rational asset pricing bubbles. *Econometrica: Journal of the Econometric Society*, 19–57.

Shi, S. (1997). A divisible search model of fiat money. *Econometrica: Journal of the Econometric Society*, 75–102.

Sidrauski, M. (1967). Inflation and economic growth. *Journal of political economy 75*(6), 796–810.

Stockman, A. C. (1981). Anticipated inflation and the capital stock in a cash in-advance economy. *journal of Monetary Economics 8*(3), 387–393.

Svensson, L. E. (2000). The zero bound in an open economy: A foolproof way of escaping from a liquidity trap. Technical report, National Bureau of Economic Research.

Tirole, J. (1985). Asset bubbles and overlapping generations. *Econometrica: Journal of the Econometric Society*, 1071–1100.

Tobin, J. (1965). Money and economic growth. *Econometrica: Journal of the Econometric Society*, 671–684.

Werning, I. (2011). Managing a liquidity trap: Monetary and fiscal policy. Technical report, National Bureau of Economic Research.

Appendix

Omitted Proofs

Proposition 1

Proof. The properties of λ and σ_c are straightforward from equation (8), noticing only that $\rho - ((1-\lambda)\sigma)^2 = r - (\hat{x} - \delta) > 0$ in equilibrium. First, for $\beta > 0$, $\lambda \in (0, 1)$ because k_t, and $m_t + h_t$ are strictly positive. The left hand side of (8) is increasing in λ while the right hand side is decreasing. Since the right hand side is increasing in σ, it follows that λ is increasing in σ. When $\sigma = 0$, (8) simplifies to $\lambda = \beta$, while in the limit as $\sigma \to \infty$, since $\lambda > 0$ we must have $(1 - \lambda) \to 0$.

The right hand side is increasing in the idiosyncratic consumption risk $\sigma_c = (1 - \lambda)\sigma$, so since λ is increasing in σ, so must σ_c. When $\sigma = 0$ we have $\lambda = \beta$, so $\sigma_c = 0$. When $\sigma \to \infty$ we have $\lambda \to 1$, so $\sigma_c = (1 - \lambda)\sigma \to \sqrt{\rho(1 - \beta)} > 0$.

For $\beta = 0$ we clearly have $\lambda = 0$ from (11). □

Proposition 2

Proof. Straightforward from equilibrium conditions $r = a - \delta - \sigma^2$ and $\hat{x} = a - \rho$. □

Proposition 3

Proof. We know from Proposition 1 that λ and $\sigma_c = (1 - \lambda)\sigma$ are increasing in σ. So the real interest rate $r = a - \delta - (1 - \lambda)\sigma^2$ falls when σ increases, but less so than without money, since in that case $r = a - \delta - \sigma^2$. Investment $\hat{x}$ falls because the term $\rho\frac{\lambda-\beta}{1-\lambda} > 0$ is increasing in λ.

If $\sigma = 0$, we have $r = a - \delta$, $\lambda = \frac{\rho\beta}{\rho} = \beta$, and $\hat{x} = a - \rho$ which coincide with the non-monetary economy. □

Proposition 4

Proof. The inflation target π does not appear in equations (14), (15), and (8) for r, $\hat{x}$, and λ. It only appears in equation (9) for $\hat{m}$.

Targeting the nominal interest rate is accomplished by picking the inflation target π so that $i = r + \pi$ is constant. Since π does not affect r, or any other real variable, we can do this to hit any

$i > 0$. Notice that targeting i introduces the usual indeterminacy in the price level, since only the expected inflation is pinned down.

Finally, since nothing real is affected by changing the inflation target, and $i \approx 0$ maximizes the utility from money services m, the optimal inflation target delivers the Friedman rule. □

Proposition 5

Proof. From individual optimization we know that the expenditure share of liquidity is β, invariant to b. Since r, $\hat{x}$, and λ are pinned down by (14), (15), and (8), they are not affected by changes in b. □

Proposition 6

Proof. It suffices to look at the behavior of λ defined by (8), which we can rewrite

$$\lambda \left(\rho - ((1-\lambda)\sigma)^2\right) = \rho\beta$$

As $\beta \to 0$ the left hand side converges to zero, and so must the right hand side. This means that either a) $\lambda \to 0$ or b) $\lambda \to 1 - \frac{\sqrt{\rho}}{\sigma}$. If $\sigma < \sqrt{\rho}$, then $1 - \frac{\sqrt{\rho}}{\sigma} < 0$. Since $\lambda > \beta$ always, b) cannot be, so we are left with a) $\lambda \to 0$. From (14) and (15) we see that r and x converge to the values on the non-monetary economy with $\lambda = \beta = 0$.

If instead $\sigma \geq \sqrt{\rho}$, we cannot have $\lambda \to 0$, because it implies that $\rho - ((1-\lambda)\sigma)^2 \leq 0$ at some point along the way (for λ small enough). Since $\lambda > \beta$ always, this requires $\rho\beta < 0$, which is not true. So we have b) $\lambda \to 1 - \frac{\sqrt{\rho}}{\sigma} \geq 0$, and the inequality is strict if $\sigma > \sqrt{\rho}$. From (14) and (15) we see that the real interest rate r is high and investment $\hat{x}$ low relative to the economy without money ($\beta = 0$). □

Proposition 7

Proof. See Online Appendix □

Proposition 8

Proof. Combine (28) with $\sigma_c = \frac{\rho(1-\beta)\sigma}{a-\hat{x}}$ to obtain

$$\sigma_c = \frac{\rho(1-\beta)}{\rho + \sigma_c^2}\sigma$$

It follows that when $\sigma = 0$ we get $\sigma_c = 0$, and σ_c is increasing in σ. Rewrite it

$$\frac{\sigma_c}{\sigma} = \frac{\rho(1-\beta)}{\rho + \left(\frac{\sigma_c}{\sigma}\right)^2 \sigma^2}$$

It follows that σ_c/σ is decreasing in σ. The properties of $\hat{x}$ follow from equation (28). Finally, write

$$r = \rho + \hat{x} - \delta - \sigma_c^2 = a - \delta - 2\sigma_c^2$$

It follows that r falls with σ. □

Proposition 9

Proof. First take the $\beta \in (0, 1/2)$ case. Rewrite (29) and (30) in terms of $y = \sigma_c/\sigma$

$$\hat{x}_{SP} = a - \rho - y^2\sigma^2 \qquad \text{Social Planner}$$
$$\hat{x}_{CE} = a - \rho - y\sigma^2 \times (1 - y) \qquad \text{Competitive Equilibrium}$$

and the incentive-compatible combinations

$$\hat{x}_{IC} = a - \frac{\rho(1-\beta)}{y}$$

The competitive equilibrium lies at the intersection of $\hat{x}_{CE}$ and $\hat{x}_{IC}$; call the corresponding $y_{CE} \in [0, 1-\beta]$. The planner's allocation lies at the intersection of $\hat{x}_{SP}$ and $\hat{x}_{IC}$, call the corresponding $y_{SP} \in [0, 1-\beta]$.

We know that $y = \sigma_c/\sigma$ can range from 0 to $1 - \beta$, in both the CE and SP (the upper bound comes from knowing that investment is below the first best in both the CE and the SP, and using $\hat{x}_{IC}$.

$\hat{x}_{IC}$ is increasing, strictly concave, and ranges from $-\infty$ when $y = 0$ to the first best $a - \rho$ when $y = 1 - \beta$. It does not depend on σ, so it will be fixed when we do comparative statics.

$\hat{x}_{CE}$ and $\hat{x}_{SP}$ do depend on σ. They both start at the first best $a - \rho$ when $y = 0$. $\hat{x}_{SP}$ is strictly decreasing and concave (it's an inverted parabola) with vertex at $(0, a - \rho)$. So it must cross $\hat{x}_{IC}$ exactly once.

$\hat{x}_{CE}$ is a parabola with vertex at $(\frac{1}{2}, a - \rho - \frac{1}{4}\sigma^2)$. Importantly, it intersects with $\hat{x}_{SP}$ at this point. For $\sigma > 0$ they intersect at exactly two points, corresponding to $y = 0$ and $y = 1/2$, and this implies that $\hat{x}_{CE} < \hat{x}_{SP}$ for all $y \in (0, 1/2)$, and $\hat{x}_{CE} > \hat{x}_{SP}$ for all $y \in (1/2, 1-\beta)$. Finally, $\hat{x}_{CE} < a - \rho$ for all $y \in (0, 1-\beta)$. In particular, $\hat{x}_{CE}(1-\beta) = a - \rho - \sigma^2\beta(1-\beta) < a - \rho$. Since $\hat{x}_{CE}$ is strictly convex and $\hat{x}_{IC}$ is strictly concave they intersect at two points at most. Since $\hat{x}_{IC} = \infty$ for $y = 0$, $\hat{x}_{IC}$ crosses $\hat{x}_{CE}$ first from below, and the from above. But since $\hat{x}_{IC}(1-\beta) = 1 - \rho > \hat{x}_{CE}(1-\beta)$, the second intersection has $y > 1 - \beta$, so it is not in the range on y. There is then only one valid intersection between $\hat{x}_{CE}$ and $\hat{x}_{IC}$; we called it $y_{CE} \in [0, 1-\beta]$ and $\hat{x}_{IC} < \hat{x}_{CE}$ for all $y < y_{CE}$ and $\hat{x}_{IC} > \hat{x}_{SP}$ for all $y > y_{CE}$ in the range of y.

Now the lower envelope of $\hat{x}_{CE}$ and $\hat{x}_{SP}$, $\hat{x}_L = \min\{\hat{x}_{CE}, \hat{x}_{SP}\}$ coincides with $\hat{x}_{CE}$ for $y \in [0, 1/2]$ and with $\hat{x}_{SP}$ for $y \in [1/2, 1-\beta]$. This implies that if $\hat{x}_{IC}$ first intersects with the lower envelope for $y < 1/2$, it must do so at y_{CE}, and if it first intersects at $y > 1/2$, it must do so at y_{SP}. In the first case, since $\hat{x}_{SP} > \hat{x}_{CE}$ for $y < 1/2$, it is strictly decreasing, and goes from $a - \rho$ for $y = 0$ to

$a - \rho - \frac{1}{4}\sigma^2 = \min \hat{x}_{CE}$ for $y = 1/2$; and $\hat{x}_{IC}$ is strictly increasing and goes to $1-\rho$; then it means that $y_{SP} < 1/2$ as well and $y_{SP} > y_{CE}$. In the second case, obviously $y_{CE} > y_{SP} > 1/2$. If it first intersects at $y = 1/2$ then $y_{CE} = y_{SP}$.

It follows immediately that if $y_{CE} < y_{SP}$ and both below $1/2$, then $\hat{x}_{CE} < \hat{x}_{SP}$ and $\sigma_c^{CE} < \sigma_c^{SP}$. On the other hand, if $y_{CE} > y_{SP}$ and both above $1/2$, then $\hat{x}_{CE} > \hat{x}_{SP}$ and $\sigma_c^{CE} > \sigma_c^{SP}$.

It only remains to see which will hold for a given σ. Since both $\hat{x}_{CE}$ and $\hat{x}_{SP}$ are decreasing for $y \in (0, 1/2)$, and $\hat{x}_{IC}$ is always increasing, it is enough to compare their values at $y = 1/2$. If $\hat{x}_{CE} = \hat{x}_{SP} \geq \hat{x}_{IC}$ at $y = 1/2$, then y_{CE} and y_{SP} are both in $[1/2, 1-\beta]$. If instead $\hat{x}_{CE} = \hat{x}_{SP} \leq \hat{x}_{IC}$ at $y = 1/2$, then y_{CE} and y_{SP} are both in $(0, 1/2)$.

$$\hat{x}_{CE} = \hat{x}_{SP} \geq \hat{x}_{IC}$$

$$\iff a - \rho - \frac{1}{4}\sigma^2 \geq a - 2\rho(1-\beta)$$

$$\iff \sigma^2 \leq \sigma^* = 2\sqrt{\rho(1-2\beta)} > 0$$

Finally, for the case $\beta \in [1/2, 1)$, $\hat{x}_{CE} < \hat{x}_{SP}$ for all $y \in [0, 1-\beta]$, regardless of σ, so $\hat{x}_{CE} < \hat{x}_{SP}$ and $\sigma_c^{CE} < \sigma_c^{SP}$. Notice that in this case the formula for $\sigma^* < 0$. □

Proposition 10

Proof. We already know that the planner's allocation is a BGP with constant $\hat{x}$ and σ_c. By setting the subsidy/tax τ^k according to (36) we ensure that r, $\hat{x}$, and $\hat{m}$ satisfy all the conditions for a BGP equilibrium.

We can check that the value of total wealth $w = k_t + m_t + h_t - \int_t^\infty e^{-rs}\tau_s^k k_s ds$ satisfies $c_t = \rho(1-\beta)w_t$, or equivalently $\sigma_c = \frac{k_t}{w_t}\sigma$. Write

$$\frac{w_t}{k_t} = 1 + \frac{\hat{m}i}{r - (\hat{x} - \delta)} + \frac{\tau^k}{r - (\hat{x} - \delta)}$$

$$\frac{w_t}{k_t} = \frac{r - (\hat{x} - \delta) + \hat{c}\frac{\beta}{1-\beta} + \tau^k}{r - (\hat{x} - \delta)} = \frac{\rho - \sigma_c^2 + \rho\beta\frac{\sigma}{\sigma_c} + 2\sigma_c^2 - \sigma_c\sigma}{\rho - \sigma_c^2}$$

$$\frac{w_t}{k_t} = \frac{\rho + \rho\beta\frac{\sigma}{\sigma_c} + \sigma_c^2 - \sigma_c\sigma}{\rho - \sigma_c^2} = \frac{\sigma}{\sigma_c}\left(\frac{\rho\frac{\sigma_c}{\sigma} + \rho\beta + \sigma_c^2(\frac{\sigma_c}{\sigma} - 1)}{\rho - \sigma_c^2}\right)$$

Use the planner's FOC (29) and the skin in the game IC constraint (31)

$$\frac{w_t}{k_t} = \frac{\sigma}{\sigma_c}\left(\frac{\rho\frac{\sigma_c}{\sigma} + \rho\beta + (a - \hat{x} - \rho)(\frac{\sigma_c}{\sigma} - 1)}{2\rho - a - \hat{x}}\right) = \frac{\sigma}{\sigma_c}\left(\frac{\rho(1+\beta) + (a - \hat{x})\frac{\rho(1-\beta) - (a - \hat{x})}{a - \hat{x}}}{2\rho - a - \hat{x}}\right)$$

$$\frac{w_t}{k_t} = \frac{\sigma}{\sigma_c}\left(\frac{\rho(1+\beta) + \rho(1-\beta) - (a - \hat{x})}{2\rho - a - \hat{x}}\right) = \frac{\sigma}{\sigma_c}\frac{2\rho - (a - \hat{x})}{2\rho - (a - \hat{x})} = \frac{\sigma}{\sigma_c}$$

The optimal inflation target implements the Friedman rule and delivers unbounded utility from real money balances (in a supremum sense). $\square$

Proposition 11

Proof. First, to implement the optimal allocation we need $r > \hat{x} - \delta$ which is equivalent to $\rho - \sigma_c^2 > 0$. We know σ_c in the optimal allocation is given by

$$\sigma_c = \frac{\rho(1-\beta)}{\rho + \sigma_c^2}\sigma$$

and σ_c is increasing in σ. For $\sigma = 0$ we have $\sigma_c = 0$ too, so $\rho - \sigma_c^2 > 0$. So we only need to ask at what σ we have $\sigma_c^2 = \rho$:

$$\sqrt{\rho} = \frac{\rho(1-\beta)}{\rho+\rho}\bar{\sigma} \implies \bar{\sigma} = \frac{2\sqrt{\rho}}{1-\beta}$$

For $\sigma \geq \bar{\sigma}$ we have $\sigma_c^2 \geq \rho$ and therefore $r \leq \hat{x} - \delta$, so the optimal allocation cannot be implemented as a competitive equilibrium with a tax on capital.

For the sign of τ^k, use (36) to get

$$\tau^k = \sigma_c(2\sigma_c - \sigma)$$

So if $\sigma_c > \frac{1}{2}\sigma$ we have $\tau^k > 0$, and if $\sigma_c < \frac{1}{2}\sigma$ we have $\tau^k < 0$. In the optimal allocation we have

$$\frac{\sigma_c}{\sigma} = \frac{\rho(1-\beta)}{\rho + \left(\frac{\sigma_c}{\sigma}\right)^2 \sigma^2}$$

So $\sigma_c/\sigma \to 1-\beta$ when $\sigma \to 0$, and σ_c/σ is decreasing in σ. So if $\beta \geq \frac{1}{2}$ we must have $\sigma_c \leq \frac{1}{2}\sigma$ and therefore $\tau^k \leq 0$, for all $\sigma \in [0, \bar{\sigma})$. If instead $\beta < \frac{1}{2}$, we have $\tau^k > 0$ for σ close to 0. We only need to find σ^* such that $\sigma_c = \frac{1}{2}\sigma^*$.

$$\frac{1}{2} = \frac{\rho(1-\beta)}{\rho + \frac{1}{4}(\sigma^*)^2} \implies \sigma^* = 2\sqrt{\rho(1-2\beta)}$$

It only remains to show that $\sigma^* \in (0, \bar{\sigma}]$. $\sigma^* > 0$ follows from $\beta < \frac{1}{2}$. Now write

$$\sigma^* = 2\sqrt{\rho(1-2\beta)} \leq 2\sqrt{\rho} \leq \frac{2\sqrt{\rho}}{1-\beta} = \bar{\sigma}$$

$\square$

Proposition 12

Proof. From the definition of h, $\square$

$$m_t + h_t = m_t + \mathbb{E}_t^Q \left[\int_t^T e^{-\int_t^s r_u du} \frac{dM_s}{p_s} \right] + \mathbb{E}_t^Q \left[e^{-\int_t^T r_u du} h_T \right]$$

use $dm_t = \frac{dM_t}{p_t} - m_t \pi_t dt$ to write

$$= m_t + \mathbb{E}_t^Q \left[\int_t^T e^{-\int_t^s r_u du} (dm_s + \pi_s m_s ds) \right] + \mathbb{E}_t^Q \left[e^{-\int_t^T r_u du} h_T \right]$$

We can write

$$e^{-\int_t^T r_u du} m_T = m_t + \int_t^T e^{-\int_t^s r_u du} (-r_s m_s ds + dm_s)$$

and plug it in to obtain

$$m_t + h_t = \mathbb{E}_t^Q \left[\int_t^T e^{-\int_t^s r_u du} ((r_s + \pi_s) m_s ds) \right] + \mathbb{E}_t^Q \left[e^{-\int_t^T r_u du} (m_T + h_T) \right]$$

Then take the limit $T \to \infty$ and use the transversality condition

$$\lim_{T \to \infty} \mathbb{E}_t^Q \left[e^{-\int_t^T r_u du} (m_T + h_T) \right] = 0$$

and the monotone convergence theorem ($r_t + \pi_t = i_t \geq 0$) to obtain

$$m_t + h_t = \mathbb{E}_t^Q \left[\int_t^\infty e^{-\int_t^s r_u du} m_s i_s ds \right]$$

Proposition 13

Proof. From (49), we plug in $\theta^{TFP} = \tilde{\sigma}^{TFP}$, and $\theta^{RS} = \frac{\sigma_\psi}{1+\psi}$, as well as $\hat{m}i = \rho\beta w/k = \rho\beta(1+\psi)$ from (47) and

$$\hat{x} + g - \delta - r = -(\rho + \frac{\mu_\psi}{1+\psi} - (\frac{\sigma}{1+\psi})^2 - (\tilde{\sigma}^{TFP})^2 - (\frac{\sigma_\psi}{1+\psi})^2)$$

Then use Ito's lemma to obtain

$$\mu_\psi = \psi' \phi(\bar{\sigma} - \sigma) + \frac{1}{2} \psi'' \sigma \nu^2$$

$$\sigma_\psi = \psi' \sqrt{\sigma} \nu$$

The ODE (51) has the μ_ψ terms together, and the $\tilde{\sigma}^{TFP}$ terms cancel out. I also simplified the terms involving σ_ψ into one term. $\square$

Proposition 14

Proof. The equations for the competitive equilibrium are a modified version of (42)-(47), taking into account that total wealth now includes the present value of taxes/subsidies that are rebated lump-sum and the tax τ^k:

$$\begin{aligned}
r_t &= \rho + (\hat{x}_t + g - \delta) + \mu_{\hat{c},t} - \sigma_{ct}^2 - (\tilde{\sigma}_{ct}^{TFP})^2 - (\tilde{\sigma}_{ct}^{RS})^2 && \text{Euler} \quad (64)\\
r_t &= a + g - \delta - \tau_t^k - \sigma_{ct}\sigma_t - \theta_t^{TFP}\tilde{\sigma}^{TFP} && \text{asset pricing} \quad (65)\\
\sigma_{ct} &= \frac{k_t}{w_t}\sigma_t = (1-\beta)\rho k_t c_t^{-1}\sigma_t && \text{idiosyncratic risk} \quad (66)\\
\tilde{\sigma}_c^{TFP} &= \theta_t^{TFP} && \text{TFP risk} \quad (67)\\
\tilde{\sigma}_{ct}^{RS} &= \theta_t^{RS} && \text{risk shock risk} \quad (68)\\
\hat{m}_t &= \beta/(1-\beta) \times (a - \hat{x}_t)/(r_t + \pi_t) && \text{money} \quad (69)
\end{aligned}$$

Since the planner's allocation satisfies (52)-(57) and the FOC (59), it satisfies also the equilibrium conditions (64)-(69). Equations (64), (67), (68), and (69) are immediate. Equation (65) follows from plugging the definition of $\hat{x}_t$ from (59), τ_t^k from (60), and $\theta_t^{TFP} = \tilde{\sigma}_{ct}^{TFP}$ into the Euler equation (53). (66) comes from the skin in the game constraint (52), using the fact that $\sigma_{Ut} = \frac{1}{\rho}\sigma_{ct}$. Finally, $\lim_{t\to\infty} \mathbb{E}^Q\left[e^{-\int_0^t r_u du} k_t \hat{c}_t\right] \implies \lim_{t\to\infty} \mathbb{E}^Q\left[e^{-\int_0^t r_u du} w_t\right] = 0$ and $i_t = r_t + \pi_t > 0$ ensure that this is in fact an equilibrium. $\square$

Proposition 15

Let $g(n; r)$ be the stationary distribution of wealth in the steady state. Market clearing in the steady state requires

$$\int n g(n; r) dn = k_{ss}(r) + m_{ss}(r)$$

where g, k_{ss}, and m_{ss} depend on the steady state real interest rate (m_{ss} depends on the nominal interest rate, but with constant money supply, inflation in the steady state is zero, so $i = r$). We must find r that solves this equation. In general this requires finding the stationary distribution of wealth generated by each r.

Optimize $(1-\beta)\log c + \beta \log m$ subject to total expenditures $x = c + mi$ to obtain an indirect utility function. We obtain the usual demand for money

$$m = \frac{\beta}{1-\beta}\frac{c}{i}$$

and an indirect utility function $\tilde{u}(x) = \log x + A(i)$, for some constant function $A(i)$. We can then re-write the budget constraint

$$dn_{it} = (n_{it}r + w(r)e_{it} - x_t)dt$$

We conclude that for a given r, we get the same distribution of wealth $g_0(n;r)$ as in a model without money and log preferences only over consumption, $\beta = 0$.

Market clearing with money requires

$$\int ng_0(n;r)dn - k_{ss}(r) = m_{ss}(r)$$

Normalizing by the steady state wage $w_{ss}(r) = f_l'(k_{ss}(r))$, we get

$$\frac{\int ng_0(n;r)dn}{w_{ss}(r)} - \frac{k_{ss}(r)}{w_{ss}(r)} = \frac{m_{ss}(r)}{w_{ss}(r)}$$

We can now compare the economy without money, $\beta = 0$, to the economy with money, $\beta > 0$. The distribution of wealth and the capital stock is the same in both cases, but with money we get a positive term on the left hand side. It's a standard result that with these preferences and technology, the right hand side is an increasing function of r,[53] so we can conclude that in the monetary economy, the interest rate is higher, and the capital stock lower, relative to the non-monetary economy.

The role of intertemporal elasticity and risk aversion

The baseline model has log preferences, which yield clean results and are quantitatively reasonable. In this Appendix I extend the baseline model to allow for EZ preferences to understand the role of intertemporal elasticity and risk aversion.

Suppose agents have recursive EZ preferences with discount ρ, risk aversion γ, and intertemporal elasticity ψ. If $\psi = 1/\gamma$ we have the standard CRRA preferences. If $\psi = \gamma = 1$ we have the baseline model with log preferences.

The equilibrium equations are now modified as follows

$$\begin{aligned} r &= \rho + (\hat{x} - \delta)/\psi - (1 + 1/\psi)(\gamma/2)\sigma_c^2 && \text{Euler equation} \\ r &= a - \delta - \gamma\sigma_c\sigma && \text{asset pricing} \\ \sigma_c &= (1 - \lambda)\sigma && \text{risk sharing} \\ \hat{m} &= \frac{\beta}{1 - \beta}\frac{a - \hat{x}}{r + \pi} && \text{money} \end{aligned}$$

The expression for the value of liquidity, λ, must be solved simultaneously with r and $\hat{x}$.

$$\lambda = \frac{\rho\beta}{\rho + (1/\psi - 1)(\hat{x} - \delta) - (1 + 1/\psi)(\gamma/2)((1 - \lambda)\sigma)^2} \tag{70}$$

[53] See for example Light (2018).

We can check that if $\psi = \gamma = 1$ we recover the equation in the baseline model.

First Best. If there is no idiosyncratic risk, $\sigma = 0$, we get closed-form expressions for r and $\hat{x}$

$$r = a - \delta$$

$$\hat{x} = (a - \delta - \rho)\psi + \delta$$

Incomplete risk sharing and no money, $\beta = 0$. The non-monetary economy also allows for closed-form expressions, because $\lambda = \beta = 0$ and $\sigma_c = \sigma$.

$$r = \underbrace{a - \delta}_{\text{first best}} - \gamma\sigma^2$$

$$\hat{x} = \underbrace{(a - \delta - \rho)\psi + \delta}_{\text{first best}} + \psi\Big[\underbrace{(1 + 1/\psi)(\gamma/2)\sigma^2}_{\text{precautionary}} - \underbrace{\gamma\sigma^2}_{\text{risk pr.}}\Big]$$

After a risk shock increases idiosyncratic risk σ, the real interest rate falls to accommodate the higher risk premium $\alpha = \gamma\sigma_c\sigma = \gamma\sigma^2$. But investment may go up or down, depending on the intertemporal elasticity ψ. If $\psi > 1$, investment falls when idiosyncratic risk σ goes up; if $\psi < 1$, investment raises. This can be understood in terms of the risk premium and precautionary motive. If $\psi > 1$, the precautionary motive is smaller than the risk premium, and the difference increases with σ ($(1 + 1/\psi)/2 < 1$). Intuitively, capital is less attractive because it is more risky, and since agents are very intertemporally elastic, they substitute towards consuming instead (accepting a big change in the growth rate of their consumption). But if $\psi < 1$, the precautionary motive dominates. Agents really want to smooth out their utility, and since they face more risk, they make it up by accumulating more capital. If $\psi = 1$, as in the baseline, the two effects cancel out and investment does not change when σ goes up.

The important variable is the intertemporal elasticity. Risk aversion, γ, just makes the idiosyncratic risk matter more. In fact, both enter jointly $\gamma\sigma^2$ in the equations. The role of intertemporal elasticity is well understood, and is the reason that the literature on time varying risk typically assumes high intertemporal elasticity, $\psi > 1$. Empirically, evidence about ψ is mixed, but $\psi = 1$ is considered a quantitatively reasonable benchmark.

Incomplete risk sharing and money, $\beta > 0$. Now let's see what happens when we add money. First, take the value of liquidity $\lambda > 0$ as given. Idiosyncratic risk sharing improves, $\sigma_c = (1 - \lambda)\sigma$, so we get

$$r = \underbrace{a - \delta}_{\text{first best}} - \gamma(1 - \lambda)\sigma^2$$

$$\hat{x} = \underbrace{(a - \delta - \rho)\psi + \delta}_{\text{first best}} + \psi\Big[\underbrace{(1 + 1/\psi)(\gamma/2)\sigma^2(1 - \lambda)^2}_{\text{precautionary}} - \underbrace{\gamma\sigma^2(1 - \lambda)}_{\text{risk pr.}}\Big]$$

Money weakens the risk premium, so the real interest rate is higher than without money. Money also weakens the precautionary motive more than the risk premium, just as in the baseline model. But since investment can go up or down with risk, depending on ψ, it is useful to decompose the effect of higher risk into the effect without money, and what money adds relative to the non-monetary economy:

$$r = \underbrace{a - \delta - \gamma\sigma^2}_{\text{non-monetary}} + \underbrace{\lambda\gamma\sigma^2}_{\Delta r} \tag{71}$$

$$\hat{x} = \underbrace{(a - \delta - \rho)\psi + \delta + \psi\big[(1 + 1/\psi)(\gamma/2)\sigma^2 - \gamma\sigma^2\big]}_{\text{non-monetary}} + \underbrace{\gamma\sigma^2\psi\left((\lambda^2 - 2\lambda)(1 + 1/\psi)/2 + \lambda\right)}_{\Delta\hat{x}} \tag{72}$$

The second terms are the effect of money on the real interest rate, Δr, and investment, $\Delta\hat{x}$, relative to the economy without money.

In general it is possible for investment in the monetary economy to be higher than in the non-monetary one. For very large ψ, $\Delta\hat{x} \approx \gamma\sigma^2\psi\lambda^2/2 > 0$. There are two forces at work. Remember that if $\psi > 1$, the risk premium dominates, so high risk σ can have a very large negative effect on investment $\hat{x}$. Money improves risk sharing and weakens the risk premium $\alpha = \gamma\sigma^2(1 - \lambda)$, so it dampens the fall in investment from this channel. It also weakens the precautionary motive relative to the risk premium, which reduces investment just like in the baseline model. The two forces work in opposite directions. In the baseline setting the direct effect of high risk in the absence of money is zero, so money must depress investment.

For $\psi \leq 1$ we can obtain a clean characterization, such that the main properties of the baseline model go through.[54] Money keeps the real interest rate from falling during downturns with high risk, and depresses investment. Risk aversion γ and idiosyncratic risk σenter together as $\gamma\sigma^2$, so all our results apply as well to increases in risk aversion.

Proposition 16. *For $\psi \leq 1$, the monetary economy has higher interest rate and lower investment than the non-monetary one; i.e. $\Delta r > 0$ and $\Delta\hat{x} < 0$. Higher $\gamma\sigma^2$ leads to higher risk premium $\alpha = \gamma\sigma^2(1 - \lambda)$, higher value of liquidity λ, and larger Δr and $|\Delta\hat{x}|$.*

The case $\psi \leq 1$ covers two very salient classes of preferences. First, CRRA preferences with risk aversion $\gamma = 1/\psi \geq 1$. This is the most common specification in macroeconomic models. In the context of models with time varying risk, it has the unappealing feature that, without money, higher risk leads to more investment.

Second, the cleanest and quantitatively salient specification has $\psi = 1$ and $\gamma > 1$. The $\gamma > 1$ can be interpreted either as high risk aversion, or as ambiguity aversion as in Barillas et al. (2009). This specification has the advantage that the equations boils down to those of the baseline model, with the only modification of replacing σ^2 by $\gamma\sigma^2$,

$$r = \underbrace{a - \delta - \gamma\sigma^2}_{\text{non-monetary}} + \underbrace{\lambda\gamma\sigma^2}_{\Delta r}$$

[54] $\psi \leq 1$ is sufficient, but not necessary.

$$\hat{x} = \underbrace{(a - \rho)}_{\text{non-monetary}} - \underbrace{\rho\frac{\lambda - \beta}{1 - \lambda}}_{\Delta\hat{x}}$$

$$\lambda = \frac{\rho\beta}{\rho - \gamma\sigma^2(1-\lambda)^2}$$

Proof of Proposition 16

Proof. Write the equation for λ, replacing r and $\hat{x}$ from (71) and (72) to obtain

$$\lambda = \frac{\rho\beta}{\rho + (1/\psi - 1)(a - \delta - \rho)\psi - \alpha\left[(1-\psi) + \psi(1 + 1/\psi)(1-\lambda)/2\right]}$$

where $\alpha = \gamma\sigma^2(1-\lambda) > 0$ is the risk premium. The denominator is strictly decreasing in α (here we use $\psi \leq 1$) and strictly increasing in λ (for $\alpha > 0$ which must be the case for $\sigma > 0$). So if there is a solution $\lambda(\alpha)$ to this equation, it is increasing in α. From $\alpha/(1-\lambda) = \gamma\sigma^2$, since the left hand side is increasing in α, it follows that α is increasing in $\gamma\sigma^2$, and therefore so is λ.

Now $\Delta r = \lambda\gamma\sigma^2 > 0$ and increasing in $\gamma\sigma^2$ is straightforward. For $\Delta\hat{x}$ we write it after some algebra

$$\Delta\hat{x} = \gamma\sigma^2\lambda\left(\lambda\frac{1+\psi}{2} - 1\right) = \alpha\lambda\frac{\lambda\frac{1+\psi}{2} - 1}{1-\lambda} < 0$$

If $\gamma\sigma^2$ increases, so does α and λ. The derivative of the last factor with respect to λ is

$$\partial_\lambda\left\{\frac{\lambda\frac{1+\psi}{2} - 1}{1-\lambda}\right\} = \frac{\frac{1+\psi}{2}(1-\lambda) + \lambda\frac{1+\psi}{2} - 1}{(1-\lambda)^2} = \frac{\frac{1+\psi}{2} - 1}{(1-\lambda)^2} \leq 0$$

(using $\psi \leq 1$ again). So $\Delta\hat{x}$ is negative and becomes even more negative. $\square$

Online Appendix: Contractual setting

In this Appendix I develop the contractual environment that yields the incomplete idiosyncratic risk sharing problem in the baseline model as the optimal contract. I also allow aggregate risk with complete risk sharing, which is the setting in the dynamic model in section 4. The setting in the baseline model is a special case with no aggregate risk.

The setting is essentially a special case of the environment in Di Tella and Sannikov (2016) with perfect misreporting ($\phi = 1$ in the terms of that paper), generalized to allow for aggregate shocks. I discuss the similarities and differences below.

Setting

The setting is as in the dynamic model in section 4. The "capital quality" shock for an agent is

$$\Delta^k_{i,t} = \sigma_t k_{i,t} dW_{i,t} + \tilde{\sigma}^{TFP} dZ^{TFP}_t \tag{73}$$

where Z^{TFP} is an aggregate TFP shock. Aggregate TFP risk $\tilde{\sigma}^{TFP}$ is constant, but idiosyncratic risk σ_t follows an autoregressive process

$$d\sigma_t = \mu_\sigma(\sigma_t)dt + \tilde{\sigma}_\sigma(\sigma_t)dZ^{RS}_t \tag{74}$$

where Z^{RS} is the aggregate risk shock. Z^{TFP} and Z^{RS} are independent Brownian motions.

There is a complete financial market with real interest rate r, nominal interest rate i, capital's excess return α, and price of aggregate shocks θ^{TFP} and θ^{RS}, all adapted to the history of aggregate shocks Z^{TFP} and Z^{RS}. Let Q be the equivalent martingale measure associated with r, θ^{TFP} and θ^{RS}, and $\tilde{Q}$ the equivalent martingale measure associated with r, θ^{TFP} , θ^{RS}, and α.[55]

The agent receives consumption c and money holdings m from the principal, and manages capital k, all contingent on the history of aggregate shocks Z^{TFP} and Z^{RS} and the agent's report of his idiosyncratic shock Y^s. The idiosyncratic shock is not observable by the principal, so the agent can misreport at rate s, such that his reports are $Y^s_t = W_t - \int_0^t \frac{s_u}{\sigma_u} du$. Furthermore, the agent has access to hidden trade that allows him to choose his consumption $\tilde{c}$, money $\tilde{m}$, capital holdings $\tilde{k}$, and to trade aggregate risk $\tilde{\sigma}^{TFP}_n$ and $\tilde{\sigma}^{RS}_n$.[56] His hidden savings n start at $n_0 = 0$ and satisfy the dynamic budget constraint

$$\begin{aligned} dn_t = {} & (n_t r_t + c_t - \tilde{c}_t + (m_t - \tilde{m}_t)i_t + (\tilde{k}_t - k_t)\alpha_t + \theta^{TFP}_t \sigma^{TFP}_{nt} + \theta^{RS}_t \tilde{\sigma}^{RS}_{nt} + k_t s_t)dt \\ & + (\tilde{k}_t - k_t)\sigma_t dW_t + \tilde{\sigma}^{TFP}_{nt} dZ^{TFP}_t + \tilde{\sigma}^{RS}_{nt} dZ^{RS} \end{aligned} \tag{75}$$

[55] That is, Q is defined by the SPD $d\xi_t/\xi_t = -r_t - \theta^{TFP}_t dZ^{TFP}_t - \theta^{RS}_t dZ^{RS}_t$ and $\tilde{Q}$ by $d\tilde{\xi}_t/\tilde{\xi}_t = -r_t - \theta^{TFP}_t dZ^{TFP}_t - \theta^{RS}_t dZ^{RS}_t - \frac{\alpha_t}{\sigma_t} dW_t$.

[56] To keep things simple, allow $\tilde{k} < 0$, but we can also restrict it to $\tilde{k} \geq 0$, as in Di Tella and Sannikov (2016). This doesn't change the optimal contract.

with solvency constraint $n_t \geq \underline{n}_t$ where $\underline{n}_t$ is the natural debt limit

$$\underline{n}_t = -\max_{s\in\mathbb{S}} \mathbb{E}_t^{\tilde{Q}} \left[\int_t^\infty e^{\int_t^u r_\tau d\tau} (c_u(Y^s) + m_u(Y^s) i_u + s_u k_u(Y^s)) du \right] \tag{76}$$

where $\mathbb{S} = \left\{ s : \mathbb{E}^{\tilde{Q}} \left[\int_0^\infty e^{\int_t^u r_\tau d\tau} |c_u(Y^s) + m_u(Y^s) i_u + s_u k_u(Y^s)| du \right] < \infty \right\}$ is the set of feasible stealing plans for a given contract. The natural debt limit $\underline{n}_t$ is the maximum amount that the agent can pay back for sure at time t. The lender is not taking any risk as long as he enforces the natural debt limit.

Lemma 1. *Assume* $|\underline{n}_0| < \infty$. *If* $n_t \geq \underline{n}_t$ *always, then* $\liminf_{t\to\infty} e^{-\int_0^t r_u du} n_t \geq 0$ *a.s.*

A contract $\mathcal{C} = (c, m, k)$ is *admissible* if $\mathbb{E}^Q \left[\int_0^\infty e^{-\int_0^t r_u du} |c_t + m_t i_t + k_t \alpha_t| \, dt \right] < \infty$. It is always optimal to implement no misreporting or hidden trade.[57] An admissible contract is *incentive compatible* if the agent chooses to report truthfully and not engage in hidden trade,

$$(c, m, k, 0, 0, 0) \in \arg\max_{\mathcal{P}} U(c, m) \quad st: \quad (75)$$

where $\mathcal{P} = (\tilde{c}, \tilde{m}, \tilde{k}, \tilde{\sigma}_n^{TFP}, \tilde{\sigma}_n^{RS}, s)$. An incentive-compatible contract is *optimal* if it minimizes the cost of delivering utility to the agent

$$J_0(u_0) = \min_{(c,m,k)\in IC} \mathbb{E}^Q \left[\int_0^\infty e^{-\int_0^t r_u du} \left(c_t + m_t i_t - k_t \alpha_t \right) dt \right]$$

$$st: \quad U(c, m) \geq u_0$$

We pin down the agent's initial utility u_0 with a free-entry condition for principals. If the agent has initial wealth w_0, he gives it to the principal in exchange for the full-commitment contract, and the principal breaks even, $J_0(u_0) = w_0$.

Incentive compatibility and optimal contract

Given contract $\mathcal{C} = (c, m, k)$, the agent's problem is to choose a misreporting and hidden trade strategy $\mathcal{P} = (\tilde{c}, \tilde{m}, \tilde{k}, \tilde{\sigma}_n^{TFP}, \tilde{\sigma}_n^{RS}, s)$ to maximize his utility subject to his dynamic budget constraint. With the natural debt limit, the dynamic budget constraint is equivalent to the following intertemporal budget constraint

$$\mathbb{E}^{\tilde{Q}} \left[\int_0^\infty e^{-\int_0^t r_u du} (\tilde{c}_t + \tilde{m}_t i_t) dt \right] \leq \max_{s\in\mathbb{S}} \mathbb{E}^{\tilde{Q}} \left[\int_0^\infty e^{-\int_0^t r_u du} (c_t(Y^s) + m_t(Y^s) i_t + k_t(Y^s) s_t) dt \right] \tag{77}$$

The right hand side is the present value of the agent's income from the principal, including what he "steals" from him, and is equal to (minus) the natural debt limit $-\underline{n}_0$. Of course, if the right hand side is infinity the agent can achieve infinite utility. This corresponds to the case where the natural debt limit $\underline{n}_0 = -\infty$ so the agent can get infinite utility under the dynamic constraint as well.

[57] See Di Tella and Sannikov (2016).

Lemma 2. *Assume $|\underline{n}_0| < \infty$. If $(\tilde{c}, \tilde{m}, \tilde{k}, \tilde{\sigma}_n^{TFP}, \tilde{\sigma}_n^{RS}, s)$ and n satisfy the dynamic budget constraint (75) with $n_t \geq \underline{n}_t$ always, then $(\tilde{c}, \tilde{m})$ satisfy the intertemporal budget constraint (77).*

If $(\tilde{c}, \tilde{m})$ satisfy the intertemporal budget constraint (77), then there are processes $(\tilde{k}, \tilde{\sigma}_n^{TFP}, \tilde{\sigma}_n^{RS}, s)$ and n that satisfy the dynamic budget constraint (75) with $n_t \geq \underline{n}_t$ always.

We can split the agent's problem into two parts. First, pick a misreporting strategy that maximizes the value of the right hand side. Second, choose $\tilde{c}$ and $\tilde{m}$ to maximize utility subject to the intertemporal budget constraint (77).

If $s^* = 0$ is optimal, then

$$\int_0^t e^{-\int_0^u r_\tau d\tau}(c_u(Y^s) + m_u(Y^s)i_u + k_u(Y^s)s_u)du - e^{-\int_0^t r_u du}\underline{n}_t(Y^s)$$

must be a $\tilde{Q}$-martingale for $s = 0$ and a supermartingale for any other s. So we can write

$$\begin{aligned} d\left(e^{-\int_0^t r_u du}\underline{n}_t(Y^s)\right) = e^{-\int_0^t r_\tau d\tau}\big\{&(c_t(Y^s) + m_t(Y^s)i_t)dt + \sigma_{\underline{n}t}(Y^s)\left(dY_t^s + \alpha_t dt\right) \\ &+\tilde{\sigma}_{\underline{n}t}^{TFP}(Y^s)(dZ_t^{TFP} + \theta_t^{TFP}dt) + \tilde{\sigma}_{\underline{n}t}^{RS}(Y^s)(dZ_t^{RS} + \theta_t^{RS}dt)\big\} \end{aligned}$$

If the agent misreports s, then

$$\begin{aligned} &\int_0^t e^{-\int_0^u r_\tau d\tau}\{(c_u(Y^s) + m_u(Y^s)i_u + k_u(Y^s)s_u)du - (c_u(Y^s) + m_u(Y^s)i_u)du \\ &-\sigma_{\underline{n}u}(Y^s)\left(dY_u^s + \alpha_u du\right) - \tilde{\sigma}_{\underline{n}u}^{TFP}(Y^s)(dZ_u^{TFP} + \theta_u^{TFP}du) - \tilde{\sigma}_{\underline{n}u}^{RS}(Y^s)(dZ_u^{RS} + \theta_u^{RS}du)\} \end{aligned}$$

or simplifying,

$$\begin{aligned} \int_0^t e^{-\int_0^u r_\tau d\tau}\{&k_u(Y^s)s_u du - \sigma_{\underline{n}u}(Y^s)\left(dY_u^s + \alpha_u du\right) - \tilde{\sigma}_{\underline{n}u}^{TFP}(Y^s)(dZ_u^{TFP} + \theta_u^{TFP}du) \\ &-\tilde{\sigma}_{\underline{n}u}^{RS}(Y^s)(dZ_u^{RS} + \theta_u^{RS}du)\} \end{aligned}$$

must be a $\tilde{Q}$-supermartingale. Since $dY_t^s = dW_t - \frac{s_t}{\sigma_t}dt$, this requires

$$k_t + \sigma_{\underline{n}t}\frac{1}{\sigma_t} = 0 \implies \sigma_{\underline{n}t} = -k_t\sigma_t$$

In other words, for every dollar the agent misreports he must lose a dollar in present value of future income.

Second, taking the right hand side of the intertemporal budget constraint as given and choosing $\tilde{c}$ and $\tilde{m}$ is a standard consumption-portfolio problem. The FOC are

$$\mu_{ct} = r_t - \rho + \sigma_{ct}^2 + (\tilde{\sigma}_{ct}^{TFP})^2 + (\tilde{\sigma}_{ct}^{RS})^2 \qquad \text{Euler equation} \qquad (78)$$

$$\alpha_t = \sigma_{ct}\sigma_t \qquad \text{demand for capital} \qquad (79)$$

$$\tilde{\sigma}_{ct}^{TFP} = \theta_t^{TFP} \qquad \text{TFP shocks} \qquad (80)$$

$$\tilde{\sigma}_{ct}^{RS} = \theta_t^{RS} \qquad \text{risk shocks} \tag{81}$$

$$m_t/c_t = \beta/(1-\beta)i_t^{-1} \qquad \text{money} \tag{82}$$

In addition, optimality and zero hidden savings, $n_t = 0$, imply that

$$\mathbb{E}_t^{\tilde{Q}}\left[\int_t^\infty e^{-\int_t^u r_\tau d\tau}(\tilde{c}_u + \tilde{m}_u i_u)du\right] = \mathbb{E}_t^{\tilde{Q}}\left[\int_t^\infty e^{-\int_t^u r_\tau d\tau}(c_u(Y^0) + m_u(Y^0)i_u)dt\right] = -\underline{n}_t$$

and $c_t = (1-\beta)\rho\mathbb{E}_t^{\tilde{Q}}\left[\int_t^\infty e^{-\int_t^u r_\tau d\tau}(\tilde{c}_u + \tilde{m}_u i_u)du\right] = (1-\beta)\rho(-\underline{n}_t)$. From this it follows that the geometric volatility of consumption is

$$\sigma_{ct} = (1-\beta)\rho\frac{k_t}{c_t}\sigma_t \qquad \text{"skin in the game"} \tag{83}$$

It's worth noting that, given (78)-(82), the agent's continuation utility

$$U_t = \mathbb{E}_t\left[\int_t^\infty e^{-\rho u}((1-\beta)\log c_u + \beta\log m_u)du\right]$$

will admit a representation $U_t = A_t + \frac{1}{\rho}\log c_t$. The skin in the game constraint (83) then implies $\sigma_{Ut} = (1-\beta)c_t^{-1}k_t\sigma_t$. If the agent misreports a dollar and immediately consumes it he gets marginal utility $(1-\beta)c_t^{-1}$, so his continuation utility must go down by that amount to deter him.

Putting these conditions together we obtain the following result.

Lemma 3. *An incentive-compatible contract $\mathcal{C} = (c, m, k)$ must satisfy conditions* (78)-(83).

The incentive compatibility conditions (78)-(83) are necessary. In general, proving that they are sufficient for global incentive compatibility is a difficult problem, because the hidden trade allows the agent a large set of deviations. The strategy is to first characterize the optimal contract subject only to the necessary incentive compatibility constraints, and then prove that it is indeed incentive compatible. As it turns out, this will be straightforward in this setting because the optimal contract will coincide with letting the agent choose his own consumption, money, and capital (the optimal contract coincides with autarky).

We say that contract $\mathcal{C} = (c, m, k)$ *solves the portfolio problem for* $w_0 > 0$ if it maximizes $U(c, m)$ subject to the dynamic budget constraint

$$\begin{aligned} dw_t = {} & (r_t w_t - c_t - m_t i_t + k_t\alpha_t + \theta_t^{TFP}\tilde{\sigma}_{wt}^{TFP} + \theta_t^{RS}\tilde{\sigma}_{wt}^{RS})dt + k_t\sigma dW_t \\ & + \tilde{\sigma}_{wt}^{TFP}dZ_t^{TFP} + \tilde{\sigma}_{wt}^{RS}dZ_t^{RS} \end{aligned} \tag{84}$$

with solvency constraint $w_t \geq 0$. This dynamic budget constraint is equivalent to

$$\mathbb{E}^{\tilde{Q}}\left[\int_0^\infty e^{-\int_0^t r_u du}(c_t + m_t i_t)dt\right] \leq w_0 \tag{85}$$

It is well known that (78)-(82) are the FOCs for this portfolio problem, so we get the following result.

Theorem 1. *Let (c, m, k) be an optimal contract for initial utility u_0, with cost $J(u_0)$. Then (c, m, k) solves the portfolio problem for $w_0 = J(u_0)$.*

Conversely, let (c, m, k) solve the portfolio problem for some $w_0 > 0$. If in addition $\lim_{t\to\infty} \mathbb{E}[e^{-rt}w_t] = 0$, then (c, m, k) is an optimal contract for initial utility u_0 with $J(u_0) = w_0$.

Remark. The condition $\lim_{t\to\infty} \mathbb{E}[e^{-rt}w_t] = 0$ must be satisfied in the competitive equilibrium in the paper.

Comparison to Di Tella and Sannikov (2016)

This setting is essentially the same as in Di Tella and Sannikov (2016), with hidden investment and perfect misreporting ($\phi = 1$ in the context of that paper). The main result here is Theorem 1, which is analogous to Lemma 28 in that paper. This is therefore a special case of the environment in that paper.

But there are some differences. First, here I allow aggregate risk shocks that affect the investment environment. The setting in Di Tella and Sannikov (2016) is stationary. Second, in Di Tella and Sannikov (2016) the agent faces a no-debt solvency constraint $n_t \geq 0$ on his hidden savings n. Here I allow the agent to borrow up to the natural borrowing limit, using his income from the contract. As it turns out the optimal contract is the same. The no-debt borrowing constraint relaxes the IC constraints, but the principal does not use this freedom in the optimal contract. Intuitively, with $n_t \geq 0$ the principal could backload the agent's consumption if he wanted. But what he really wants to do is to front load it.

Finally, here I allow the agent to short capital in his hidden investment, $\tilde{k}_t < 0$ and to overreport returns, $s_t < 0$. This is done for simplicity. In Di Tella and Sannikov (2016) hidden investment and misreporting must be non-negative, $k_t \geq 0$ and $s_t \geq 0$, and the optimal contract is the same (for the special case with $\phi = 1$).

Proofs

Proof of Lemma 1

Proof. From the definition of the natural debt limit (76), if we take absolute value on both sides we get the following inequality

$$|\underline{n}_t| \leq S_t = \mathbb{E}_t^{\tilde{Q}} \left[\int_t^\infty e^{\int_t^u r_\tau d\tau} |c_u(Y^{s^*}) + m_u(Y^{s^*}) i_u + s_u^* k_u(Y^{s^*})| du \right] < \infty$$

where s^* is the misreporting process that achieves the maximum in (76). The martingale representation theorem yields

$$d\left(e^{-\int_0^t r_u du} S_t\right) = -e^{-\int_0^t r_u du} |c_t(Y^{s^*}) + m_t(Y^{s^*}) i_t + s_t^* k_t(Y^{s^*})| dt + \tilde{Q}\text{-local mart. terms}$$

We also know that $\lim_{T\to\infty} \mathbb{E}^{\tilde{Q}}\left[e^{-\int_0^T r_u du} S_T\right] = 0$. To see this, write

$$S_0 = \mathbb{E}_0^{\tilde{Q}}\left[\int_0^T e^{\int_0^u r_\tau d\tau} |c_u(Y^{s^*}) + m_u(Y^{s^*})i_u + s_u^* k_u(Y^{s^*})| du\right] + \mathbb{E}^{\tilde{Q}}\left[e^{-\int_0^T r_u du} S_T\right]$$

and take the limit $T \to \infty$, using the MCT on the first term. It follows that $\lim_{T\to\infty} e^{-\int_0^T r_u du} S_T$ exists and is zero almost surely (see Problem 3.16 in Karatzas and Shreve (2012)). Since $|\underline{n}_t| \leq S_t$, the same is true for $\underline{n}_t$, and since $n_t \geq \underline{n}_t$, we obtain $\liminf_{T\to\infty} e^{-\int_0^t r_u du} n_t \geq 0$ a.s. □

Proof of Lemma 2

Proof. In the first direction, use the dynamic budget constraint to compute

$$\begin{aligned}
\mathbb{E}^{\tilde{Q}}\left[e^{-\int_0^t r_u du} n_t\right] = {} & \mathbb{E}^{\tilde{Q}}\left[\int_0^t e^{-\int_0^u r_\tau d\tau}(c_u(Y^s) + m_u(Y^s)i_u + k_u(Y^s)s_u) du\right] \\
& - \mathbb{E}^{\tilde{Q}}\left[\int_0^t e^{-\int_0^u r_\tau d\tau}(\tilde{c}_u + \tilde{m}_u i_u) dt\right]
\end{aligned}$$

Subtract $\mathbb{E}^{\tilde{Q}}\left[e^{-\int_0^t r_u du} \underline{n}_t\right] < \infty$ from both sides. Because $\underline{n}_0$ is the maximum value that the agent can get, we obtain an inequality:

$$\begin{aligned}
\mathbb{E}^{\tilde{Q}}\left[e^{-\int_0^t r_u du}(n_t - \underline{n}_t)\right] \leq {} & \max_s \mathbb{E}^{\tilde{Q}}\left[\int_0^\infty e^{-\int_0^u r_\tau d\tau}(c_u(Y^s) + m_u(Y^s)i_u + k_u(Y^s)s_u) du\right] \\
& - \mathbb{E}^{\tilde{Q}}\left[\int_0^t e^{-\int_0^u r_\tau d\tau}(\tilde{c}_u + \tilde{m}_u i_u) du\right]
\end{aligned}$$

$$\mathbb{E}^{\tilde{Q}}\left[\int_0^t e^{-\int_0^u r_\tau d\tau}(\tilde{c}_u + \tilde{m}_u i_u) du\right] \leq -\underline{n}_0 - \mathbb{E}^{\tilde{Q}}\left[e^{-\int_0^t r_u du}(n_t - \underline{n}_t)\right]$$

Take the limit $t \to \infty$ and use $n_t \geq \underline{n}_t$ to obtain the intertemporal budget constraint (77).

In the other direction, define

$$n_t = \underline{n}_t + \mathbb{E}^{\tilde{Q}}\left[\int_t^\infty e^{-\int_t^u r_\tau d\tau}(\tilde{c}_u + \tilde{m}_u i_u) du\right] \geq \underline{n}_t$$

Define $L_t = \mathbb{E}_t^{\tilde{Q}}\left[\int_t^\infty e^{-\int_t^u r_\tau d\tau}(\tilde{c}_u + \tilde{m}_u i_u) du\right]$, so that $\int_0^t e^{-\int_0^u r_\tau d\tau}(\tilde{c}_u + \tilde{m}_u i_u) du + e^{-\int_0^t r_\tau d\tau} L_t$ is a $\tilde{Q}$-martingale. Likewise, $-\int_0^t e^{-\int_0^u r_\tau d\tau}(c_u(Y^{s^*}) + m_u(Y^{s^*})i_u + k_u(Y^{s^*})s_u^*) dt + e^{-\int_0^t r_\tau d\tau} \underline{n}_t$ is also $\tilde{Q}$-martingale, where s^* is the misreporting process that achieves the maximum. So we can write

$$\begin{aligned}
dn_t = {} & \left(n_t r_t + c_t(Y^{s^*}) + m_t(Y^{s^*})i_t + k_t(Y^{s^*})s_t^* - (\tilde{c}_t + \tilde{m}_t i_t)\right) dt \\
& + (\sigma_{\underline{n}t} + \sigma_{Lt})(\alpha_t dt + dW_t) + (\tilde{\sigma}_{\underline{n}t}^{TFP} + \tilde{\sigma}_{Lt}^{TFP})(\theta_t^{TFP} dt + d\tilde{Z}_t^{TFP}) + (\tilde{\sigma}_{\underline{n}t}^{RS} + \tilde{\sigma}_{Lt}^{RS})(\theta_t^{RS} dt + d\tilde{Z}_t^{RS})
\end{aligned}$$

Letting $\sigma_{\underline{n}t} + \sigma_{Lt} = \tilde{k}_t - k_t$, $\tilde{\sigma}_{\underline{n}t}^{TFP} + \tilde{\sigma}_{Lt}^{TFP} = \tilde{\sigma}_{nt}^{TFP}$, and $\tilde{\sigma}_{\underline{n}t}^{RS} + \tilde{\sigma}_{Lt}^{RS} = \tilde{\sigma}_{nt}^{RS}$, we obtain the dynamic budget constraint (75). □

Proof of Lemma 3

Proof. Immediate from the argument in section 6, noting that incentive compatibility requires $|\underline{n}_0| < \infty$. $\square$

Proof of Theorem 1

Proof. In the first direction, if (c, m, k) is an optimal contract, then it must satisfy the local IC constraints (78)-(82), which are the FOC for the consumption-portfolio problem. So c and m solve the optimal portfolio problem for some initial w_0, with an associated wealth process w that satisfies the dynamic budget constraint (84) and $w_t \geq 0$. Now the IC constraint (83) pins down the corresponding k. We know that $c_t = (1-\beta)\rho w_t$ in the portfolio problem, so (83) and (79) imply

$$\sigma_c = \frac{\alpha_t}{\sigma_t} = (1-\beta)\rho \frac{k_t}{c_t}\sigma_t$$

$$\implies \frac{k_t}{w_t} = \frac{\alpha_t}{\sigma_t^2}$$

which is the expression for capital in the portfolio problem. Finally, we need to show that $w_0 = J_0$. Integrate the dynamic budget constraint (84) and take expectations under Q to obtain

$$w_0 = \mathbb{E}^Q\left[\int_0^T e^{-rt}(c_t + m_t i - k_t\alpha)dt\right] + \mathbb{E}^Q\left[e^{-rT}w_T\right]$$

If we take the limit $T \to \infty$, the first term will converge to $\mathbb{E}\left[\int_0^\infty e^{-rt}(c_t + m_t i - k_t\alpha)dt\right] = J_0$ (apply dominated converge theorem and use feasibility). For the second term, because everything is proportional to w, we must have $J_t = Aw_t$ for some $A > 0$ (it will be $A = 1$). The continuation cost of the contract $J_t = \mathbb{E}_t\left[\int_t^\infty e^{-rs}(c_s + m_s i - \alpha k_s)ds\right]$ must satisfy $\lim_{T\to\infty}\mathbb{E}\left[e^{-rT}J_T\right] = 0$, so $\lim_{T\to\infty}\mathbb{E}[e^{-rT}w_T] = 0$ and therefore $w_0 = J_0$. To see why $\lim_{T\to\infty}\mathbb{E}\left[e^{-rT}J_T\right] = 0$, write $J_0 = \mathbb{E}\left[\int_0^T e^{-rt}(c_t + m_t i - k_t\alpha)dt\right] + \mathbb{E}\left[e^{-rT}J_T\right]$ and take $T \to \infty$ (using DCT and feasibility again).

In the other direction, suppose (c, m, k) solve the portfolio problem with associated wealth process w and utility utility $U(w) = \frac{1}{\rho}(\log w - \log\xi)$. Notice that this is the only contract that satisfies the local IC constraints and delivers utility $u_0 = \frac{1}{\rho}(\log w_0 - \log\xi)$. (78)-(82) are the FOC for the portfolio problem, and pin down c and m up to an initial constant (corresponding to w_0). We also know that $k_t/w_t = \frac{\alpha_t}{\sigma_t^2}$ and $c_t = (1-\beta)\rho w_t$, so from (79) we get the skin in the game IC constraint (83)

$$\sigma_{ct} = \frac{\alpha_t}{\sigma_t} = \frac{k_t}{w_t}\sigma_t = (1-\beta)\rho\frac{k_t}{c_t}\sigma_t$$

The contract (c, m, k) is feasible because $\lim_{T\to\infty}\mathbb{E}\left[e^{-rT}w_T\right] = 0$ (each term is proportional to w, so it grows slower than r). The contract is globally incentive compatible, because the agent is only getting risk-free debt from the principal, and doing what he wants.

It only remains to show that we can't achieve more utility. Integrate the dynamic budget constraint to obtain

$$w_0 = \mathbb{E}\left[\int_0^T e^{-rt}(c_t + m_t i - k_t \alpha)dt\right] + \mathbb{E}\left[e^{-rT} w_T\right]$$

and take the limit $T \to \infty$ to obtain $J_0 = w_0 > 0$. So giving the agent more utility (scaling up the contract) will cost more. $\square$