On Money as a Medium of Exchange in Near-Cashless Credit Economies

Ricardo Lagos*
New York University

Shengxing Zhang†
London School of Economics

March 2018

Abstract

We study the transmission of monetary policy in credit economies where money serves as a medium of exchange. We find that—in contrast to current conventional wisdom in policy-oriented research in monetary economics—the role of money in transactions can be a powerful conduit to asset prices and ultimately, aggregate consumption, investment, output, and welfare. Theoretically, we show that the cashless limit of the monetary equilibrium (as the cash-and-credit economy converges to a pure-credit economy) need not converge to the equilibrium of the nonmonetary pure-credit economy. Quantitatively, we find that the magnitudes of the responses of prices and allocations to monetary policy in the monetary economy are sizeable—even in the cashless limit. Hence, as tools to assess the effects of monetary policy, monetary models without money are generically poor approximations—even to idealized highly developed credit economies that are able to accommodate a large volume of transactions with arbitrarily small aggregate real money balances.

Keywords: Asset prices; Collateral; Credit; Leverage; Liquidity; Margin; Monetary policy
JEL classification: D83, E52, G12

*Lagos is thankful for the support from the C.V. Starr Center for Applied Economics at NYU.

†Zhang is thankful for the support from the Centre for Macroeconomics at LSE and the British Academy/Leverhulme Small Research Grant.

1 Introduction

A large body of work in macroeconomics rests on the premise that artificial economies without money are well suited to study the effects of monetary policy. In fact, most of the work in modern monetary economics that caters to policymakers, abstracts from the usefulness of money altogether: there is typically no money in the models, or if there is money, it is merely held as a redundant asset. What underlies this moneyless approach to monetary economics is the received wisdom that the medium-of-exchange role of money is quantitatively irrelevant in the transmission of monetary policy.

The intuitive argument runs as follows: aggregate real money balances are a small fraction of aggregate real output in modern economies (e.g., inverse velocity of the monetary base tends to be relatively small), so policy induced changes in real money balances are bound to have very small effects on output. Therefore, the argument goes, the traditional monetarist mechanism whereby changes in monetary policy are transmitted to the economy through changes in real money balances is basically irrelevant, and there is no significant loss in basing monetary policy advise on models where real money balances do not interact with the real allocation—or are simply assumed to be equal to zero.

This intuition has been formalized in the context of economies where the role of money in exchange is not modeled explicitly, but rather, is proxied by either assuming money is an argument of a utility function, or by imposing that certain purchases be paid for with cash acquired in advance. There are two ways in which these reduced-form models have been used to support the claim that medium-of-exchange considerations can be safely ignored.

First, the fact that the monetary equilibrium in these reduced-form models is continuous under a certain "cashless limit" (e.g., obtained by taking to zero either the marginal utility of real balances in a model where money enters the utility function, or the fraction of "cash goods" in a cash-credit goods version of a cash-in-advance model where money is a redundant asset) has been used to conclude that a monetary economy with an inverse velocity that is as small as in the data can be well approximated by an economy where real money balances are simply assumed to play no role in monetary transmission—an economy without money, for instance. Second, parametrized versions of these reduced-form models have been used to claim that for realistic values of inverse velocity, the quantitative relevance of real balances in monetary transmission is insignificant.

In this paper we show that when the trading frictions that make money useful in exchange are modeled explicitly, the medium-of-exchange role of money is a significant and resilient channel for the transmission of monetary policy. Specifically, we show that the two formal arguments that have been put forward to justify and encourage the use of models without money to study monetary policy, are overturned when we replace the reduced-form formulations with explicit more primitive micro foundations. First, we show that the cashless limit of the prices and allocations in the monetary equilibrium (as the cash-and-credit economy converges to a pure-credit economy) need not converge to the prices and allocations of the economy without money. So as a matter of pure theory, it would be incorrect to regard the economy with no money as an arbitrarily good approximation to an economy where credit has developed sufficiently to render equilibrium aggregate real balances negligible. Second, we show that this discontinuity is quantitatively significant: the effects of monetary policy in the explicit medium-of-exchange economy remain large even as aggregate real balances converge to zero along the cashless limit. The key insight is that along the cashless limit, even as the volume of transactions financed with cash converges to zero, individual investors always have the option to trade with cash, so the market price for these vanishing cash transactions feeds back into the terms of trade negotiated in all credit transactions (e.g., through the terms of trade for the inside loans that coexist with money). We show that this logic applies as long as credit markets are not perfectly frictionless, in the sense that credit-market intermediaries have some market power, and there exists some limit to how much an individual can borrow.

An important mechanism for our results for near-cashless economies is the fact that real balances influence the terms of trade in financial markets. The fact that a trader's asset holdings affect the terms of trade in a bilateral bargain is commonplace in models of decentralized trade. The mechanism arises naturally in models of OTC trade with unrestricted asset holdings such as Afonso and Lagos (2015). In the search-based monetary literature there are also environments where money confers a strategic bargaining advantage to the agent who holds it. In Zhu and Wallace (2007), for example, the mechanism is embedded in the bargaining protocol, according to which holding money is akin to having more bargaining power. A more recent example is Rocheteau et al. (2018), where holding money improves a borrower's outside option in the bilateral bargain for a loan. Since in their model this outside option is an increasing function of the borrower's real balances, this mechanism offers a theory for the passthrough from the nominal policy interest rate to the real borrowing rate that the money holder has to pay for the loan. Some of the trading situations that agents encounter in our model exhibit a similar mechanism, but it turns out that these particular trading situations are not the relevant ones for our main results. What prevents the medium-of-exchange transmission mechanism from dissipating in near-cashless economies is the fact that money affects the terms of trade in some transactions that *do not involve money*. The reason is that *the option to engage in monetary trades* can improve the bargaining position of an investor when he negotiates with a financial intermediary, even though neither the investor nor the intermediary holds money nor wishes to hold money. The basic structure of our model builds on Lagos and Wright (2005). The particular marketstructure is similar to the one we have used in Lagos and Zhang (2015, 2017), which in turn adopts some elements from Duffie et al. (2005). The major difference with Lagos and Zhang (2015, 2017) is that here we allow investors to buy assets on margin. Aside from capturing an important aspect of trade in financial markets, credit is essential to study the role of monetary policy in near-cashless economies, which is our main objective in this paper.

The rest of the paper is organized as follows. Section 2 presents the basic model. Equilibrium is characterized in Section 3. Section 4 presents the theoretical results for the cashless limit. Section 5 develops two extensions. The first illustrates the relevance of the medium-of-exchange transmission mechanism of monetary policy for aggregate variables such as consumption and investment. The second explores the robustness of the main results to alternative credit arrangements. Section 6 studies efficiency and welfare. Section 7 proves anlaytical results for the effects of monetary policy on asset prices and real macro aggregates. Section 8 conducts quantitative theoretical exercises designed to gauge the magnitude and empirical relevance of the medium-of-exchange transmission mechanism. Section 9 places our contribution in the context of the existing literature on monetary economics without money. Section 10 concludes. The appendix contains all proofs.

2 Model

2.1 Environment

Time is represented by a sequence of periods indexed by $t = 0, 1, \ldots$. Each time period is divided into two subperiods where different activities take place. There is a continuum of infinitely lived *investors*, each identified with a point in the set $\mathcal{I} = [0, N_I]$, with $N_I \in \mathbb{R}_+$. There is also a continuum of infinitely lived *brokers* of two types, denoted B and E. Each broker of type B, or *bond broker*, is identified with a point in the set $\mathcal{B} = [0, N_B]$, with $N_B \in \mathbb{R}_+$. Each broker of type E, or *equity broker*, is identified with a point in the set $\mathcal{E} = [0, N_E]$, with $N_E \in \mathbb{R}_+$. All agents discount payoffs across periods with the same factor, $\beta \in (0, 1)$.

There is a continuum of production units with measure $A^s \in \mathbb{R}_{++}$ that are active every period. Every active unit yields an exogenous *dividend* $y_t \in \mathbb{R}_+$ of a perishable consumption good at the end of the first subperiod of period t. (Each active unit yields the same dividend as every other active unit, so $y_t A^s$ is the aggregate dividend.) At the beginning of every period, every active unit is subject to an independent idiosyncratic shock that renders it permanently unproductive with probability $1 - \eta \in [0, 1)$. If a production unit remains active, its dividend in period t is $y_t = \gamma_t y_{t-1}$ where γ_t is a nonnegative random variable with cumulative distribution function Γ, i.e., $\Pr(\gamma_t \leq \gamma) = \Gamma(\gamma)$, and mean $\bar{\gamma} \in (0, (\beta\eta)^{-1})$. The time t dividend becomes known to all agents at the beginning of period t, and at that time each failed production unit is replaced by a new unit that yields dividend y_t in the initial period and follows the same stochastic process as other active units thereafter (the dividend of the initial set of production units, $y_0 \in \mathbb{R}_{++}$, is given at $t = 0$). In the second subperiod of every period, every agent has access to a linear production technology that transforms the agent's effort into a perishable homogeneous consumption good (the *general good*).

For each active production unit there is a durable and perfectly divisible *equity share* that represents the bearer's ownership of the production unit and confers the right to collect dividends. At the beginning of every period $t \geq 1$, each investor receives an endowment of $(1 - \eta) A^s$ equity shares corresponding to the new production units. (When a production unit fails, its equity share disappears.) There is a second financial instrument, money, that is intrinsically useless (it is not an argument of any utility or production function, and unlike equity, ownership of money does not constitute a right to collect any resources). The quantity of money at time t is denoted A_t^m. The initial quantity of money, $A_0^m \in \mathbb{R}_{++}$, is given and $A_{t+1}^m = \mu A_t^m$, with $\mu \in \mathbb{R}_{++}$. A monetary authority injects or withdraws money via lump-sum transfers or taxes to investors in the second subperiod of every period. At the beginning of period $t = 0$, each investor is endowed with a portfolio of equity shares and money.

As is standard in monetary theory, in order to preserve a meaninful role for money, we assume investors are anonymous and unable to commit. If in addition there were also complete lack of enforcement, investors would be unable to borrow, and would have no alternative but to fund first-subperiod purchases with money. In order to allow for credit in purchases of equity, we incorporate a limited form of enforcement by letting investors issue bonds that are collateralized by the equity shares they own. Specifically, some investors can issue bonds in the first subperiod of t, each representing a claim to one unit of the general good to be delivered in the second subperiod of t. We assume the bond issued in the first subperiod of t is collateralized in the sense that if the debtor defaults when the bond is due (at the beginning of the second subperiod of t), then the creditor (i.e., the bond holder) appropriates a fraction $\lambda \in [0, 1]$ of the equity shares that the debtor owns at the time of default.[1] All financial instruments (equity shares, money, private bonds) are perfectly recognizable and cannot be forged.

The market structure is as follows. In the second subperiod, all agents can trade the consumption good produced in that subperiod, equity shares, and money in a spot Walrasian market. In the first subperiod, brokers and investors trade equity shares, money, and collateralized bonds. Trading in the first subperiod is organized as follows. Equity brokers have access to a Walrasian *equity market* where they can trade equity and money. Bond brokers have access to a Walrasian *bond market* where they can trade collateralized bonds and money. Investors access the bond and equity markets indirectly, by engaging in bilateral trades with brokers whom they meet at random. Specifically, let χ_k^j be an indicator function that equals 1 if investor $j \in \mathcal{I}$ has a trading opportunity with a broker of type $k \in \{E, B\}$ in the first subperiod. Then $\alpha_{ks} \equiv \Pr(\chi_E^j = k, \chi_B^j = s)$ for $k, s \in \{0, 1\}$, with $\sum_{k,s \in \{0,1\}} \alpha_{ks} = 1$.[2] Once a broker and an investor have contacted each other, the pair negotiates the quantities of equity shares and money (if the broker is an equity broker), or the quantities of bonds and money (if the broker is a bond broker), that the broker will trade in the corresponding first-subperiod competitive market on behalf of the investor, and an intermediation fee for the broker's services. We assume the terms of the trade between an investor and a broker are determined by Nash bargaining, where an investor has bargaining power $\theta \in [0, 1]$ in negotiations with bond brokers, and all the bargaining power in negotiations with equity brokers.[3] The timing is that the round of trades

[1] This credit arrangement is similar to the one in Barro (1976) and Kiyotaki and Moore (1997, 2005).

[2] Below, we will refer to an investor $j \in \mathcal{I}$ with $(\chi_E^j, \chi_B^j) = (k, s)$ as an *investor (of type) ks*. Notice our assumptions imply: (a) investors of type 10 and equity brokers can only trade equity and money, (b) investors of type 01 and bond brokers can only trade bonds and money, and (c) investors of type 11 can trade bonds, equity, and money.

[3] The assumption that investors have all the bargaining power when negotiating with equity brokers is made for analytical simplicity. It amounts to regarding the equity market as a conventional Walrasian market where investors bear zero transaction costs, which is a good approximation to organized exchanges such as the New York Stock Exchange. We have extensively studied the role of search and bargaining frictions in equity markets in previous work (see, e.g., Lagos and Zhang 2015, 2017), so here we focus on the role that these frictions play in the credit market.

in the first-subperiod ends before production units yield dividends. Hence equity is traded *cum dividend* in the first subperiod and *ex dividend* in the second subperiod.[4] Figure **??** illustrates the timeline and market stucture.

An individual broker's preferences are represented by

$$\mathbb{E}_0^j \sum_{t=0}^{\infty} \beta^t (c_t - h_t),$$

where $j \in \{E, B\}$ denotes the broker's type, c_t is consumption of the homogeneous good that is produced, traded, and consumed in the second subperiod of period t, and h_t is the utility cost from exerting h_t units of effort to produce this good. The expectation operator $\mathbb{E}_0^j$ is with respect to the probability measure induced by the dividend process and the random trading process in the first subperiod. Brokers get no utility from the dividend good.[5] An individual investor's preferences are represented by

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(\varepsilon_t y_t + c_t - h_t \right),$$

where y_t is the quantity of the dividend good that an investor consumes at the end of the first subperiod of period t, c_t is consumption of the homogeneous good that is produced, traded, and consumed in the second subperiod of period t, and h_t is the utility cost from exerting h_t units of effort to produce this good. The variable ε_t denotes the realization of a valuation shock that is distributed independently over time and across investors, with a differentiable cumulative distribution function G on the support $[\varepsilon_L, \varepsilon_H] \subseteq [0, \infty]$, and $\bar{\varepsilon} = \int \varepsilon dG(\varepsilon)$. Each investor learns the realization ε_t at the beginning of period t. The expectation operator $\mathbb{E}_0$ is with respect to the probability measure induced by the dividend process, the investor's valuation shock, and the random trading process in the first subperiod.

2.2 Discussion of institutional background and modeling assumptions

In modern financial markets, brokers and dealers allow certain investors to purchase eligible securities *on margin*. "Margin" is an extension of credit from a broker-dealer to an investor

[4]As in conventional search models of financial over-the-counter markets, e.g., see Duffie et al. (2005) and Lagos and Rocheteau (2009), an investor must own the equity share in order to consume the dividend.

[5]This assumption implies that brokers have no direct consumption motive for holding the equity share. It is easy to relax, but we adopt it because it is the standard benchmark in the search-based literature on over-the-counter markets, see, e.g., Duffie et al. (2005), Lagos and Rocheteau (2009), Lagos, Rocheteau, and Weill (2011), and Weill (2007).

using the investor's own securities as collateral. Funds borrowed on margin may be used for any purpose, including the purchase of securities. In particular, funds borrowed on margin are often used for the purchase of the same securities being pledged as collateral. Interest on the borrowed funds accrues over the period of time that the loan is outstanding. The use of margin is regulated by financial regulatory organizations, certain securities exchanges, and the broker-dealer holding the margin account. Buying on margin is standard practice among sophisticated investors. Hedge funds, for example, typically work with *prime brokers* that offer a package of services, including securities lending, global custody, and financing in the form of margin loans.[6] Retail brokerage companies offer small or less sophisticated investors cash brokerage accounts that require all securities purchases to be funded with cash. However, they often also offer investors who meet certain additional requirements, margin accounts that allow to borrow against eligible securities.[7]

The typical margin loan works as follows. A broker extends an investor a loan of L dollars in order to purchase A dollars worth of an asset, e.g., a stock. The investor's own capital or *equity* in this transaction is $E = A - L$. The stock is pledged as collateral to secure the loan, and E is also known as the *haircut* (or *downpayment*). The investor's equity, E, expressed as a proportion of the value of the stock, A, is the *margin* on the loan, i.e., $\mathcal{M} = E/A$. In other words, the *margin* is the proportion of the value of the purchase of a security financed by the investor's own funds. The investor's *leverage* is the reciprocal of the margin, i.e., $\mathcal{L} = A/E$, and the *loan-to-value ratio* is $\mathcal{R} = L/A$.

Margin accounts are subject to the rules of the Federal Reserve Board, the Financial Industry Regulatory Authority (FINRA), and securities exchanges such as the New York Stock Exchange, as well as the brokerage firm's margin policies, which are sometimes more stringent than those of the regulators. The Federal Reserve Board Regulation T that regulates the extension of credit by securities brokers and dealers in the United States, specifies a minimum *initial margin requirement* of 50% for new, or initial, stock purchases. (The Federal Reserve has the authority to change this initial margin requirement, but it has been constant since 1974.) The rules of FINRA and the exchanges supplement the requirements of Regulation T by placing additional *maintenance margin requirements* on customer accounts. In particular, FINRA Rule 4210

[6] Prime brokers are large investment banks or securities firms. The top list includes Bank of America Meryll Lynch, Barclays, Credit Suisse, Deutsche Bank, Goldman Sacks, J.P. Morgan, Morgan Stanley, and UBS.

[7] Examples of retail brokers that offer margin accounts to small investors include Charles Schwab Corporation, E-Trade, Fidelity Investments, Scottrade, TD Ameritrade, TradeStation, and USAA Brokerage.

requires that a customer maintains a minimum margin of 25% at all times. Under these rules, if the customer's margin falls below the minimum maintenance margin, which may be set by some brokers to be higher than 25%, the customer may be required to deposit more funds in the margin account to meet the minimum maintenance margin requirement. This is referred to as a *margin call.* Failure to meet the margin call may cause the broker to liquidate the securities in the customer's account in order to bring the account back up to the required maintenance margin level.

The loans that investors receive (by issuing collateralized bonds) in the first subperiod, are the theoretical counterparts of margin loans in actual financial markets. Suppose, as is the case in the model, that the investor can borrow up to a fraction λ of the value of the stock, and chooses to borrow that amount, i.e., $L = \lambda A$. In this case, the margin is $\mathcal{M} = 1 - \lambda$, leverage is $\mathcal{L} = (1 - \lambda)^{-1}$, and the loan-to-value ratio is $\mathcal{R} = \lambda$. In the model, the investor has no need to roll over the margin loan, so $\mathcal{M} = 1 - \lambda$ can be interpreted as the *initial margin requirement*, which is constant and typically similar across brokerage firms and eligible stocks.[8] In the model, as in actual markets, investors who want to buy stocks and are able to obtain margin loans from brokers can take on leverage, while those who cannot get margin loans must rely on their own funds.

We model margin loans as bilateral agreements with terms that are negotiated between the parties, as is the case for large investors in real-world markets. In the model, an investor may trade equity with no access to a margin loan (with probability α_{10}), may lend funds and not trade equity (with probability α_{01}), or may trade equity with access to margin loans (with probability α_{11}). This is a flexible and simple way to capture the multitude of circumstances investors may face in actual financial markets. For example, $\alpha_{10} > 0$ captures the fact that on a given day some investors may be unable to secure (mutually acceptable terms for) margin loans from brokers. The market structure in the first subperiod represents a prototypical over-the-counter (OTC) trading arrangement that involves finding a suitable counterparty and then negotiating the terms of the trade. Notice this general stylized OTC structure nests several perfectly competitive benchmarks as special cases. For example, $\alpha_{10} = 1$ corresponds to an equity market that is competitive and frictionless (i.e., with no search or bargaining frictions), where all trades must be paid for with money. Alternatively, $\alpha_{11} = \theta = 1$ corresponds to

[8]Since in the theory the investor repays the loan at the end of the period when it was issued, there is no role for maintenance margin requirements or margin calls.

competitive, frictionless, and fully integrated equity and loan markets.

2.3 Bargaining and portfolio problems

Three bargaining situations may arise in the OTC round of trade. With probability α_{10}, an investor only contacts an equity broker, and they bargain over the quantities of equity and money the investor buys or sells, and the broker's intermediation fee. In this case the bargaining protocol is that the investor makes a take-it-or-leave-it offer to the equity broker. With probability α_{01}, an investor only contacts a bond broker, and they bargain over the quantities of bonds and money the investor buys or sells, and the broker's intermediation fee (the outcome is determined by Nash bargaining with investor bargaining power θ). With probability α_{11}, an investor contacts both an equity broker and a bond broker, and simultaneously bargains with both over the quantities of equity, bonds, and money, as well as the respective intermediation fees. In this case the bargaining protocol is that the investor makes a take-it-or-leave-it offer to the equity broker, and the outcome of the negotiation with the bond broker is determined by Nash bargaining with investor bargaining power θ. Throughout we assume a broker's fee is expressed in terms of the general good and paid by the investor in the second subperiod.[9] Next, we formulate the portfolio problems faced by each type of broker at the end of the first subperiod (after the round of OTC trades), and then turn to the three bargaining situations an investor may face in the OTC round of trade.

An individual agent's portfolio at the beginning of period t is represented by a vector $\boldsymbol{a}_t = (a_t^m, a_t^s) \in \mathbb{R}_+^2$, i.e., it consits of $a_t^m \in \mathbb{R}_+$ units of money and $a_t^s \in \mathbb{R}_+$ equity shares.[10] Let $\hat{W}_t^E(\boldsymbol{a}_t, k_t)$ denote the maximum expected discounted payoff of an equity broker with portfolio $\boldsymbol{a}_t = (a_t^m, a_t^s)$ and earned fee k_t, as he reallocates his portfolio in the equity market at the end of the first subperiod of period t (i.e., after the OTC round of transactions). Let $W_t^E(\boldsymbol{a}_t, k_t)$ denote the maximum expected discounted payoff, at the beginning of the second subperiod of period t, of an equity broker who has earned fee k_t in the OTC round of period t, and is holding

[9] In related work (Lagos and Zhang, 2015), we instead assume that the investor must pay the intermediation fee on the spot, i.e., with money or equity. The alternative formulation we use here makes the analysis and the exposition much simpler while the main economic mechanisms are essentially unchanged.

[10] Since bonds issued in the OTC round of period $t-1$ are settled in the second subperiod of $t-1$, there are no bonds outstanding at the beginning of period t.

portfolio $\boldsymbol{a}_t = (a_t^m, a_t^s)$.[11] Then

$$\hat{W}_t^E(\boldsymbol{a}_t, k_t) = \max_{(\overline{a}_t^m, \overline{a}_t^s) \in \mathbb{R}_+^2} W_t^E(\overline{a}_t^m, \overline{a}_t^s, k_t) \tag{1}$$
$$\text{s.t. } \overline{a}_t^m + p_t \overline{a}_t^s \leq a_t^m + p_t a_t^s,$$

where $\boldsymbol{a}_t = (a_t^m, a_t^s)$ and p_t denotes the nominal price of an equity share in the first-subperiod equity market. We use $\bar{\boldsymbol{a}}_{Et} = (\overline{a}_{Et}^m(\boldsymbol{a}_t), \overline{a}_{Et}^s(\boldsymbol{a}_t))$ to denote the solution to (1), and refer to it as the equity broker's *post-trade portfolio*.

Let $\hat{W}_t^B(\boldsymbol{a}_t, k_t)$ denote the maximum expected discounted payoff of a bond broker with portfolio $\boldsymbol{a}_t = (a_t^m, a_t^s)$ and earned fee k_t, as he reallocates his portfolio in the bond market at the end of the first subperiod of period t (i.e., after the OTC round of transactions).[12] Let $W_t^B(\boldsymbol{a}_t, a_t^b, k_t)$ denote the maximum expected discounted payoff, at the beginning of the second subperiod of period t, of a bond broker who has earned fee k_t in the OTC round of period t, and is holding portfolio $(\boldsymbol{a}_t, a_t^b) = (a_t^m, a_t^s, a_t^b) \in \mathbb{R}_+^2 \times \mathbb{R}$ of money, equity, and bonds. Then

$$\hat{W}_t^B(\boldsymbol{a}_t, k_t) = \max_{(\overline{a}_t^m, \overline{a}_t^b) \in \mathbb{R}_+ \times \mathbb{R}} W_t^B(\overline{a}_t^m, a_t^s, \overline{a}_t^b, k_t) \tag{2}$$
$$\text{s.t. } \overline{a}_t^m + q_t \overline{a}_t^b \leq a_t^m$$
$$W_t^B(\overline{a}_t^m, (1-\lambda) a_t^s, 0, k_t) \leq W_t^B(\overline{a}_t^m, a_t^s, \overline{a}_t^b, k_t),$$

where $\boldsymbol{a}_t = (a_t^m, a_t^s)$ and q_t is the nominal price of a bond. The first constraint is the budget constraint the bond broker faces in the bond market of the OTC round. The second constraint is the collateral constraint that ensures the bond broker will prefer to repay in the following subperiod any debt he may have issued in the previous OTC round, rather than default and forfeit a fraction λ of his equity holding. Notice that the bond broker's equity holding, a_t^s, is fixed in the portfolio problem (2), since he has no access to the equity market and is therefore unable to change his equity holding in the OTC round. We use $\bar{\boldsymbol{a}}_{Bt}(\boldsymbol{a}_t) = (\overline{a}_{Bt}^m(\boldsymbol{a}_t), \overline{a}_{Bt}^s(\boldsymbol{a}_t), \overline{a}_{Bt}^b(\boldsymbol{a}_t))$, subject to $\overline{a}_{Bt}^s(\boldsymbol{a}_t) = a_t^s$, to denote the solution to (2), and refer to it as the bond broker's *post-trade portfolio*.

[11] Equity brokers' bond position is always zero since they have no access to the bond market in the OTC round (and therefore cannot trade bonds in the current period), and there are no bonds outstanding at the beginning of the period.

[12] In principle, the bond broker may be holding a nonzero bond position when reallocating his portfolio at the end of the OTC round. However, as will become clear when we formulate the relevant bargaining problem, it is without loss of generality to assume that the broker's portfolio after having provided intermediation services to an investor is the same as the broker's beginning-of-period portfolio, which has zero bonds.

Let $W_t^I(a_t^m, a_t^s, a^b, k)$ denote the maximum expected discounted payoff at the beginning of the second subperiod of period t of an investor who is holding portfolio $\left(a_t^m, a_t^s, a_t^b\right)$ and has to pay a fee k_t. Consider an investor who enters period t with (pre-trade) portfolio $\boldsymbol{a}_t = (a_t^m, a_t^s)$ and valuation ε. This investor may find himself in three bargaining situations. With probability α_{10} the investor only contacts an equity broker, and the investor's post-trade portofolio is $\overline{\boldsymbol{a}}_{10t}(\boldsymbol{a}_t, \varepsilon) = (\overline{a}_{10t}^m(\boldsymbol{a}_t, \varepsilon), \overline{a}_{10t}^s(\boldsymbol{a}_t, \varepsilon), \overline{a}_{10t}^b(\boldsymbol{a}_t, \varepsilon))$, with $\overline{a}_{10t}^b(\boldsymbol{a}_t, \varepsilon) = 0$ and

$$(\overline{a}_{10t}^m(\boldsymbol{a}_t, \varepsilon), \overline{a}_{10t}^s(\boldsymbol{a}_t, \varepsilon)) = \arg \max_{(\overline{a}_t^m, \overline{a}_t^s) \in \mathbb{R}_+^2} \varepsilon y_t \overline{a}_t^s + W_t^I\left(\overline{a}_t^m, \overline{a}_t^s, \mathbf{0}\right) \tag{3}$$
$$\text{s.t. } \overline{a}_t^m + p_t \overline{a}_t^s \leq a_t^m + p_t a_t^s,$$

where $\mathbf{0}$ is used to represent $(0, 0)$. In this case the investor is making a take-it-or-leave-it offer to the equity broker, so the bargaining outcome consists of an intermediation fee equal to zero (that leaves the broker indifferent between trading or not), and an investor post-trade portfolio given by (3) that maximizes the investor's continuation payoff. Notice that in this case the investor's bond holding is fixed at its beginning-of-period value of 0 since he has no access to the bond market and therefore cannot change his bond holding in the OTC round.

With probability α_{01} the investor only contacts a bond broker, and the bargaining outcome consists of an intermediation fee, $k_{01t}(\boldsymbol{a}_t, \varepsilon)$, and a post-trade portfolio for the investor, $\overline{\boldsymbol{a}}_{01t}(\boldsymbol{a}_t, \varepsilon) = (\overline{a}_{01t}^m(\boldsymbol{a}_t, \varepsilon), \overline{a}_{01t}^s(\boldsymbol{a}_t, \varepsilon), \overline{a}_{01t}^b(\boldsymbol{a}_t, \varepsilon))$, with $\overline{a}_{01t}^s(\boldsymbol{a}_t, \varepsilon) = a_t^s$, that solves

$$\max_{(\overline{a}_t^m, k_t, \overline{a}_t^b) \in \mathbb{R}_+^2 \times \mathbb{R}} \left[W_t^I(\overline{a}_t^m, a_t^s, \overline{a}_t^b, k_t) - W_t^I\left(a_t^m, a_t^s, \mathbf{0}\right)\right]^\theta k_t^{1-\theta} \tag{4}$$
$$\text{s.t. } \overline{a}_t^m + q_t \overline{a}_t^b \leq a_t^m$$
$$W_t^I\left(a_t^m, a_t^s, \mathbf{0}\right) \leq W_t^I(\overline{a}_t^m, a_t^s, \overline{a}_t^b, k_t)$$
$$W_t^I\left(\overline{a}_t^m, (1-\lambda) a_t^s, 0, k_t\right) \leq W_t^I(\overline{a}_t^m, a_t^s, \overline{a}_t^b, k_t).$$

The objective function in (4) is the standard geometric average of the investor's and the broker's gains from trade. The first constraint is the budget constraint the investor faces in the bond market of the OTC round. The second constraint ensures the trade is incentive compatible for the investor and the restriction $k_t \in \mathbb{R}_+$ ensures the trade is incentive compatible for the bond broker. The third constraint is the collateral constraint that ensures the investor will prefer to repay in the following subperiod any debt he may have issued in the previous OTC round, rather than default and forfeit a fraction λ of his equity holding. Notice that the investor's equity holding, a_t^s, is fixed in the bargaining/portfolio problem (4), since the investor has no access to the equity market and therefore cannot change his equity holding in the OTC round.

With probability α_{11} the investor can simultaneously trade with an equity broker and a bond broker. In this case the bargaining outcome consists of an intermediation fee to the bond broker, $k_{11t}(\boldsymbol{a}_t,\varepsilon)$, and a post-trade portfolio, $\overline{\boldsymbol{a}}_{11t}(\boldsymbol{a}_t,\varepsilon) = (\overline{a}^m_{11t}(\boldsymbol{a}_t,\varepsilon), \overline{a}^s_{11t}(\boldsymbol{a}_t,\varepsilon), \overline{a}^b_{11t}(\boldsymbol{a}_t,\varepsilon))$, that solves

$$\max_{(\overline{\boldsymbol{a}}_t,k_t)\in\mathbb{R}^2_+\times\mathbb{R}\times\mathbb{R}_+} \left[\varepsilon y_t \overline{a}^s_t + W^I_t(\overline{\boldsymbol{a}}_t,k_t) - \varepsilon y_t \overline{a}^s_{10t}(\boldsymbol{a}_t,\varepsilon) - W^I_t\left(\overline{a}^m_{10t}(\boldsymbol{a}_t,\varepsilon), \overline{a}^s_{10t}(\boldsymbol{a}_t,\varepsilon), \mathbf{0}\right)\right]^\theta k_t^{1-\theta} \quad (5)$$

$$\text{s.t. } \overline{a}^m_t + p_t\overline{a}^s_t + q_t\overline{a}^b_t \leq a^m_t + p_t a^s_t$$

$$\varepsilon y_t \overline{a}^s_{10t}(\boldsymbol{a}_t,\varepsilon) + W^I_t\left(\overline{a}^m_{10t}(\boldsymbol{a}_t,\varepsilon), \overline{a}^s_{10t}(\boldsymbol{a}_t,\varepsilon), \mathbf{0}\right) \leq \varepsilon y_t \overline{a}^s_t + W^I_t(\overline{\boldsymbol{a}}_t,k_t)$$

$$W^I_t\left(\overline{a}^m_t, (1-\lambda)\,\overline{a}^s_t, 0, k_t\right) \leq W^I_t(\overline{\boldsymbol{a}}_t,k_t),$$

where $\overline{\boldsymbol{a}}_t = \left(\overline{a}^m_t, \overline{a}^s_t, \overline{a}^b_t\right)$. Recall that in this situation the bargaining protocol is that the investor makes a take-it-or-leave-it offer to the equity broker. This has two implications. First, the equity broker's gain from trade, i.e., her intermediation fee, equals zero (just as in the bargaining solution (3)). Second, if the investor and the bond broker where unable to reach an agreement, the investor can still trade with the equity broker at the terms specified by (3). Hence the outcome (3) is the investor's outside option in his bargaining problem with the bond broker. Therefore the investor's gain from trade corresponding to an outcome $(\overline{\boldsymbol{a}}_t,k_t)$ from the joint negotiation with both brokers, consists of the continuation payoff $\varepsilon y_t\overline{a}^s_t + W^I_t(\overline{\boldsymbol{a}}_t,k_t)$, minus the investor's outside option, $\varepsilon y_t\overline{a}^s_{10t}(\boldsymbol{a}_t,\varepsilon) + W^I_t\left(\overline{a}^m_{10t}(\boldsymbol{a}_t,\varepsilon), \overline{a}^s_{10t}(\boldsymbol{a}_t,\varepsilon), \mathbf{0}\right)$, namely the payoff the investor achieves in (3). The first constraint is the budget constraint the investor faces in the OTC round when he is able to trade simultaneously in the equity and the bond market. The second constraint ensures the trade is incentive compatible for the investor (the restriction $k_t \in \mathbb{R}_+$ ensures the trade is also incentive compatible for the bond broker, and the equity broker is willing to participate since her fee is equal to 0). The third constraint is the collateral constraint that ensures the investor will prefer to repay in the following subperiod any debt he may have issued in the previous OTC round, rather than default and forfeit a fraction λ of his equity holding.

Let $V^j_t(\boldsymbol{a}_t)$ denote the maximum expected discounted payoff of a broker of type $j \in \{E, B\}$ who enters the OTC round of period t with portfolio $\boldsymbol{a}_t \equiv (a^m_t, a^s_t)$. Define $\boldsymbol{\phi}_t \equiv (\phi^m_t, \phi^s_t)$, where ϕ^m_t denotes the real price of money, and ϕ^s_t the real *ex dividend* price of equity in the second subperiod of period t (both expressed in terms of the second subperiod consumption good). Then,

$$W^E_t(\boldsymbol{a}_t) = \max_{(c_t,h_t,\tilde{\boldsymbol{a}}_{t+1})\in\mathbb{R}^4_+} \left[c_t - h_t + \beta\mathbb{E}_t V^E_{t+1}(\boldsymbol{a}_{t+1})\right] \quad (6)$$

$$\text{s.t. } c_t + \boldsymbol{\phi}_t \tilde{\boldsymbol{a}}_{t+1} \leq h_t + \boldsymbol{\phi}_t \boldsymbol{a}_t,$$

where $\tilde{\boldsymbol{a}}_{t+1} \equiv \left(\tilde{a}^m_{t+1}, \tilde{a}^s_{t+1}\right)$, $\boldsymbol{a}_{t+1} = \left(\tilde{a}^m_{t+1}, \eta \tilde{a}^s_{t+1}\right)$, $\mathbb{E}_t$ is the conditional expectation over the next-period realization of the dividend, and $\boldsymbol{\phi}_t \boldsymbol{a}_t$ denotes the dot product of $\boldsymbol{\phi}_t$ and $\boldsymbol{a}_t$. Similarly,

$$W^B_t(\boldsymbol{a}_t, a^b_t, k_t) = \max_{(c_t, h_t, \tilde{\boldsymbol{a}}_{t+1}) \in \mathbb{R}^4_+} \left[c_t - h_t + \beta \mathbb{E}_t V^B_{t+1}\left(\boldsymbol{a}_{t+1}\right)\right] \tag{7}$$

$$\text{s.t. } c_t + \boldsymbol{\phi}_t \tilde{\boldsymbol{a}}_{t+1} \leq h_t + \boldsymbol{\phi}_t \boldsymbol{a}_t + a^b_t + k_t.$$

Let $V^I_t\left(\boldsymbol{a}_t, \varepsilon\right)$ denote the maximum expected discounted payoff of an investor with valuation ε and portfolio $\boldsymbol{a}_t \equiv (a^m_t, a^s_t)$ at the beginning of the OTC round of period t. Then

$$W^I_t(\boldsymbol{a}_t, a^b_t, k_t) = \max_{(c_t, h_t, \tilde{\boldsymbol{a}}_{t+1}) \in \mathbb{R}^4_+} \left[c_t - h_t + \beta \mathbb{E}_t \int V^I_{t+1}\left(\boldsymbol{a}_{t+1}, \varepsilon\right) dG(\varepsilon)\right] \tag{8}$$

$$\text{s.t. } c_t + \boldsymbol{\phi}_t \tilde{\boldsymbol{a}}_{t+1} \leq h_t + \boldsymbol{\phi}_t \boldsymbol{a}_t + a^b_t - k_t + T_t,$$

where $\boldsymbol{a}_{t+1} = (\tilde{a}^m_{t+1}, \eta \tilde{a}^s_{t+1} + (1-\eta) A^s)$ and $T_t \in \mathbb{R}$ is the real value of the time t lump-sum monetary transfer.

The value function of an equity broker who enters the OTC round of t with portfolio $\boldsymbol{a}_t$ is

$$V^E_t\left(\boldsymbol{a}_t\right) = W^E_t\left[\bar{\boldsymbol{a}}_{Et}\left(\boldsymbol{a}_t\right)\right]. \tag{9}$$

This value function already embeds the fact that equity brokers get no gains from bilateral trades with investors in the OTC round. The value function of a bond broker who enters the OTC round of period t with portfolio $\boldsymbol{a}_t$ is

$$\begin{aligned} V^B_t\left(\boldsymbol{a}_t\right) = {} & \left(1 - \alpha^B_{01} - \alpha^B_{11}\right) W^B_t\left[\bar{\boldsymbol{a}}_{Bt}\left(\boldsymbol{a}_t\right), 0\right] \\ & + \alpha^B_{01} \int W^B_t\left[\bar{\boldsymbol{a}}_{Bt}\left(\boldsymbol{a}_t\right), k_{01t}\left(\tilde{\boldsymbol{a}}_t, \varepsilon\right)\right] dH_{It}\left(\tilde{\boldsymbol{a}}_t, \varepsilon\right) \\ & + \alpha^B_{11} \int W^B_t\left[\bar{\boldsymbol{a}}_{Bt}\left(\boldsymbol{a}_t\right), k_{11t}\left(\tilde{\boldsymbol{a}}_t, \varepsilon\right)\right] dH_{It}\left(\tilde{\boldsymbol{a}}_t, \varepsilon\right), \end{aligned} \tag{10}$$

where H_{It} is the joint cumulative distribution function over the portfolios and valuations of the investors the bond broker may contact in the OTC market of period t, α^B_{01} is the probability a bond broker contacts an investor who does not contact an equity broker, and α^B_{11} is the probability a bond broker contacts an investor who simultaneously contacts an equity broker. The value function of an investor who enters the OTC round of period t with portfolio $\boldsymbol{a}_t$ and valuation ε is

$$\begin{aligned} V_t^I(\boldsymbol{a}_t,\varepsilon) &= \alpha_{00}\left[\varepsilon y_t a_t^s + W_t^I(\boldsymbol{a}_t,\mathbf{0})\right] \\ &+ \alpha_{10}\left[\varepsilon y_t \overline{a}_{10t}^s(\boldsymbol{a}_t,\varepsilon) + W_t^I(\overline{\boldsymbol{a}}_{10t}(\boldsymbol{a}_t,\varepsilon),0)\right] \\ &+ \alpha_{01}\left[\varepsilon y_t a_t^s + W_t^I(\overline{\boldsymbol{a}}_{01t}(\boldsymbol{a}_t,\varepsilon),k_{01t}(\boldsymbol{a}_t,\varepsilon))\right] \\ &+ \alpha_{11}\left[\varepsilon y_t \overline{a}_{11t}^s(\boldsymbol{a}_t,\varepsilon) + W_t^I(\overline{\boldsymbol{a}}_{11t}(\boldsymbol{a}_t,\varepsilon),k_{11t}(\boldsymbol{a}_t,\varepsilon))\right]. \end{aligned} \tag{11}$$

3 Equilibrium

Let A_{jt}^m and A_{jt}^s denote the quantities of money and equity shares, respectively, held by all agents of type $j \in \{B,E,I\}$, i.e., bond brokers, equity brokers, and investors, at the beginning of the OTC round of period t (after production units have depreciated and been replaced). That is, $A_{jt}^m = N_j \int a_t^m dF_{jt}(\boldsymbol{a}_t)$ and $A_{jt}^s = N_j \int a_t^s dF_{jt}(\boldsymbol{a}_t)$, where F_{jt} is the cumulative distribution function over portfolios $\boldsymbol{a}_t = (a_t^m, a_t^s)$ held by agents of type j at the beginning of the OTC round of period t. Let $\tilde{A}_{jt+1}^m$ and $\tilde{A}_{jt+1}^s$ denote the total quantities of money and shares held by all agents of type $j \in \{B,E,I\}$ at the end of period t, i.e., $\tilde{A}_{Bt+1}^k = \int_{\mathcal{B}} \tilde{a}_{bt+1}^k db$, $\tilde{A}_{Et+1}^k = \int_{\mathcal{E}} \tilde{a}_{et+1}^k de$ and $\tilde{A}_{It+1}^k = \int_{\mathcal{I}} \tilde{a}_{it+1}^k di$ for $k \in \{s,m\}$, where e.g., $\tilde{a}_{bt+1}^k$ denotes the quantity of asset k held at the end of period t by the individual bond broker identified with the point $b \in \mathcal{B}$. Then $A_{jt+1}^m = \tilde{A}_{jt+1}^m$ and $A_{jt+1}^s = \eta \tilde{A}_{jt+1}^s$ for $j \in \{B,E\}$, and $A_{It+1}^m = \tilde{A}_{It+1}^m$ and $A_{It+1}^s = \eta \tilde{A}_{It+1}^s + (1-\eta) A^s$. Let $\bar{A}_{jt}^m$, $\bar{A}_{jt}^s$, and $\bar{A}_{jt}^b$ denote the quantities of money, shares, and bonds, respectively, held after the OTC round o f trade of period t by all the brokers of type $j \in \{B,E\}$. Recall that the random matching in the OTC round partitions the set of investors into four types depending on their trading opportunities: those who only contact an equity broker (labeled "10"), those who only contact a bond broker (labeled "01"), those who contact both and equity broker and a bond broker (labeled "11"), and those who contact neither (labeled "00"). Let $\bar{A}_{hnt}^m$ and $\bar{A}_{hnt}^s$ denote the quantities of money and shares held after the OTC round of trade of period t by all the investors of type hn, for $h,n \in \{0,1\}$. For asset $k \in \{s,m,b\}$, $\bar{A}_{jt}^k = N_j \int \bar{a}_{jt}^k(\boldsymbol{a}_t) dF_{jt}(\boldsymbol{a}_t)$ for $j \in \{B,E\}$. For asset $k \in \{s,m,b\}$ and investor type hn, with $h,n \in \{0,1\}$, $\bar{A}_{hnt}^k = \alpha_{hn} N_I \int \overline{a}_{hnt}^k(\boldsymbol{a}_t,\varepsilon) dH_{It}(\boldsymbol{a}_t,\varepsilon)$, where $\overline{a}_{00t}^m(\boldsymbol{a}_t,\varepsilon) = a_t^m$, $\overline{a}_{00t}^s(\boldsymbol{a}_t,\varepsilon) = a_t^s$, and $\overline{a}_{00t}^b(\boldsymbol{a}_t,\varepsilon) = 0$. We are now ready to define equilibrium.

Definition 1 *An equilibrium is a sequence of prices,* $\{p_t, q_t, \phi_t^m, \phi_t^s\}_{t=0}^{\infty}$*, portfolio allocations and fees in the OTC market,* $\{\bar{\boldsymbol{a}}_{Bt}(\cdot), \bar{\boldsymbol{a}}_{Et}(\cdot), \{\bar{\boldsymbol{a}}_{jkt}(\cdot), k_{jkt}(\cdot)\}_{j,k\in\{0,1\}}\}_{t=0}^{\infty}$*, and end-of-day* *portfolios,* $\{\tilde{\boldsymbol{a}}_{Bt+1}, \tilde{\boldsymbol{a}}_{Et+1}, \tilde{\boldsymbol{a}}_{It+1}\}_{t=0}^{\infty}$*, such that for all* t*: (i) the portfolios and fees in the OTC market solve (1), (2), (3), (4), (5); (ii) taking prices and the bargaining protocol as given, the end-of period portfolios solve (6), (7), (8); and (iii) prices are such that all Walrasian markets clear, i.e.,* $\tilde{A}^s_{Bt+1} + \tilde{A}^s_{Et+1} + \tilde{A}^s_{It+1} = A^s$ *(the end-of-period* t *Walrasian market for equity clears),* $\tilde{A}^m_{Bt+1} + \tilde{A}^m_{Et+1} + \tilde{A}^m_{It+1} = A^m_{t+1}$ *(the end-of-period* t *Walrasian market for money clears),* $\bar{A}^b_{Bt} + \bar{A}^b_{01t} + \bar{A}^b_{11t} = 0$ *(the period* t *OTC interdealer market for bonds clears),* $\bar{A}^s_{Et} + \bar{A}^s_{10t} + \bar{A}^s_{11t} = A^s_{Et} + (\alpha_{10} + \alpha_{11}) A^s_{It}$ *(the period* t *OTC interdealer market for equity clears), and* $[\bar{A}^m_{Bt} + \bar{A}^m_{Et} + \sum_{j,k\in\{0,1\}} \bar{A}^m_{jkt} - A^m_t]\mathbb{I}_{\{\phi^m_t > 0\}} = 0$ *(the money market clears in the OTC round of trade). An equilibrium is "monetary" if* $\phi^m_t > 0$ *for all* t *and "nonmonetary" otherwise.*

The first step toward characterizing equilibrium is to find the bargaining outcomes. For any $(y, z) \in \mathbb{R}^2$, it is convenient to define the "mixed indicator function" $\chi : \mathbb{R}^2 \to [0, 1]$ by

$$\chi(y, z) \begin{cases} = 1 & \text{if } y < z \\ \in [0, 1] & \text{if } y = z \\ = 0 & \text{if } z < y. \end{cases} \tag{12}$$

The following lemma characterizes equilibrium post-trade portfolios in the OTC market for an economy with no money.

Lemma 1 *Consider the economy with no money, and let*

$$\varepsilon^n_t \equiv \frac{\bar{\phi}^s_t - \phi^s_t}{y_t}, \tag{13}$$

where $\bar{\phi}^s_t$ *denotes the price of an equity share expressed in terms of bonds.*

(i) Consider an investor who enters the OTC round of period t *with equity holding* a^s_t *and valuation* ε*. Then:*

(a) If the investor is only able to contact an equity broker, the post-trade equity holding is $(\overline{a}^s_{10t}(a^s_t), \overline{a}^b_{10t}(a^s_t)) = (a^s_t, 0)$.

(b) If the investor is only able to contact a bond broker, the broker earns no intermediation fee, and the post-trade portfolio is $(\overline{a}^s_{01t}(a^s_t), \overline{a}^b_{01t}(a^s_t)) = (a^s_t, 0)$.

(c) If the investor is able to contact both an equity and a bond broker, the bargaining problem has a solution only if

$$\lambda < \frac{\bar{\phi}^s_t}{\phi^s_t}, \tag{14}$$

the post-trade portfolio is

$$\overline{a}^s_{11t}\left(a^s_t,\varepsilon\right) = \chi\left(\varepsilon^n_t,\varepsilon\right)\frac{\bar{\phi}^s_t}{\bar{\phi}^s_t - \lambda\phi^s_t}a^s_t \tag{15}$$

$$\overline{a}^b_{11t}\left(a^s_t,\varepsilon\right) = \bar{\phi}^s_t\left[1 - \chi\left(\varepsilon^n_t,\varepsilon\right)\frac{\bar{\phi}^s_t}{\bar{\phi}^s_t - \lambda\phi^s_t}\right]a^s_t, \tag{16}$$

and the intermediation fee for the bond broker is

$$k_{11t}\left(a^s_t,\varepsilon\right) = \left(1-\theta\right)\left(\varepsilon - \varepsilon^n_t\right)y_t\left[\chi\left(\varepsilon^n_t,\varepsilon\right)\frac{\bar{\phi}^s_t}{\bar{\phi}^s_t - \lambda\phi^s_t} - 1\right]a^s_t. \tag{17}$$

(ii) The post-trade portfolio of a bond broker who enters the OTC round of period t with equity holding a^s_t, is $\left(\overline{a}^s_{Bt}(a^s_t), \overline{a}^b_{Bt}(a^s_t)\right) = (a^s_t, 0)$.

(iii) The post-trade portfolio of an equity broker who enters the OTC round of period t with equity holding a^s_t is $\overline{a}^s_{Et}(a^s_t) = a^s_t$.

Parts $(i)(a)$ and $(i)(b)$ of Lemma 1 state that investors who do not simultaneously contact an equity and a bond broker, are unable to trade in the OTC round.[13] Part $(i)(c)$ states that an investor who is simultaneously in contact with an equity and a bond broker can buy or sell equity and take a long or short position in bonds. Specifically, from (15) and (16), if $\varepsilon < \varepsilon^n_t$, then the investor sells all his equity for bonds. Conversely, if $\varepsilon^n_t < \varepsilon$, the investor shorts the bond in order to take a long position in equity. Condition (17) indicates the bond broker earns a fee on these transactions as long as $\varepsilon \neq \varepsilon^n_t$ (the equity broker never earns a fee according to our bargaining protocol). Parts (ii) and (iii) state that brokers only intermediate trades and do not trade on their own account in the OTC round.

Notice the nonmonetary benchmark is capable of supporting trade in the OTC round, but only among investors who contact both an equity and a bond broker. The reason is that even though only equity can be traded in the equity market, and only bonds can be traded in the bond market, investors with access to both markets can trade in those markets *simultaneously*, which implies these investors are actually able to exchange these securities at a price of $\bar{\phi}^s_t$ bonds

[13]In a nonmonetary economy, investors who only contact an equity broker can only trade equity and therefore cannot buy equity because they have no way to pay, and cannot sell equity because they have no way to get paid (also, these investors cannot sell or buy bonds because they are not in contact with the bond market). Investors who only contact a bond broker have no benefit from selling bonds and no way to buy bonds in a nonmonetary economy (since they cannot trade equity because they are not in contact with the equity market).

per equity share.[14] Since each bond is a claim to 1 unit of the second-supberiod consumption good, we can think of $\bar{\phi}_t^s$ as the real cum dividend price in the OTC round of an equity share, expressed in terms of the second-subperiod consumption good. Conversely, in the nonmonetary economy there is an implied real interest rate on bonds (expressed in terms of equity shares), denoted i_t^n. In the OTC round, an investor of type 11 can use 1 unit of equity to purchase $\bar{\phi}_t^s$ bonds. These bonds deliver $\bar{\phi}_t^s$ general goods in the following subperiod, when the relative price of general goods in terms of equity shares is $1/\phi_t^s$. Thus $i_t^n \equiv \bar{\phi}_t^s/\phi_t^s - 1$, or equivalently,

$$i_t^n = \frac{\varepsilon_t^n y_t}{\phi_t^s}. \tag{18}$$

Finally, to see why (14) is necessary for the bargaining outcome to be well defined, consider the budget constraint and the collateral constraint of an investor who contacts an equity and a bond broker, namely $\bar{\phi}_t^s \overline{a}_t^s + \overline{a}_t^b = \bar{\phi}_t^s a_t^s$, and $-\lambda \phi_t^s \overline{a}_t^s \leq \overline{a}_t^b$. These two conditions imply the borrowing constraint $-\lambda \phi_t^s a_t^s \leq \left(1 - \lambda \phi_t^s / \bar{\phi}_t^s\right) \overline{a}_t^b$. This constraint would be slack for all $\overline{a}_t^b < 0$ if (14) were violated, meaning that an investor with $\varepsilon > \varepsilon_t^n$ would be able (and willing) to take an infinitely long position in the stock. Intuitively, notice that if (14) is violated, then an investor who starts with no wealth can sell b collateralized bonds to purchase $b/\bar{\phi}_t^s$ equity shares and this leveraged purchase would leave the investor's borrowing constraint slack, since $b < \lambda \phi_t^s b / \bar{\phi}_t^s$. Notice that with (18), (14) can be written as $\lambda < 1 + i_t^n$, so (14) could only be violated if the net real interest were negative, which as we show below, will not be the case in equilibrium.

[14] Our market structure implies agents with access to the equity market (i.e., investors of type 10, investors of type 11, and equity brokers) can trade equity *and* money, while agents with access to the bond market (i.e., investors of type 01, investors of type 11, and bond brokers) can trade bonds *and* money. This market structure is different from a conventional cash-in-advance formulation, e.g., in the spirit of Lucas (1980), where agents with access to the equity market would be restricted to trading equity *for* money, and agents with access to the bond market would be restricted to trading bonds *for* money. The fact that the nonmonetary equilibrium supports some trade in equity and bonds makes it clear that our formulation does not require agents to hold cash (in advance or otherwise) to be able to trade equity or bonds. In contrast, absent money, trade in equity or bonds would be impossible in a conventional cash-in-advance formulation, which by assumption would require every equity and bond purchase to be paid in cash. Neither does the payment structure in our model fit the predetermined cash-good/credit-good taxonomy assumed in Lucas and Stokey's (1983) extension of the pure cash-in-advance model. In our model, at a given point in time, equity may resemble a "cash good" to investors who only contact an equity broker, but a "credit good" to investors who contact an equity and a bond broker simultaneously. Similarly, over time, for a given investor, equity may resemble a "cash good" at times when the investor only contacts an equity broker, or a "credit good" at times when the investor contacts an equity and a bond broker simultaneously. Another way in which our model is different from a cash-in-advance formulation is that, as shown below, transaction velocity is endogenous (it changes with the monetary policy stance), and can exceed 1 in a model period.

The following lemma characterizes equilibrium post-trade portfolios in the OTC market for an economy with money.

Lemma 2 *Consider the economy with money, and let*

$$\varepsilon_{10t}^{*} \equiv \left(p_t \phi_t^m - \phi_t^s\right) \frac{1}{y_t} \tag{19}$$

$$\varepsilon_{11t}^{*} \equiv \varepsilon_{10t}^{*} + \left(1 - q_t \phi_t^m\right) \left[\mathbb{I}_{\{q_t\phi_t^m<1\}} \frac{p_t}{q_t} + \mathbb{I}_{\{1<q_t\phi_t^m\}} \lambda \phi_t^s\right] \frac{1}{y_t}. \tag{20}$$

(i) Consider an investor who enters the OTC round of period t with portfolio $\boldsymbol{a}_t$ and valuation ε. Then:

(a) If the investor is able to contact only an equity broker, the post-trade portfolio is

$$\bar{a}_{10t}^m(\boldsymbol{a}_t, \varepsilon) = \left[1 - \chi\left(\varepsilon_{10t}^*, \varepsilon\right)\right] \left(a_t^m + p_t a_t^s\right) \tag{21}$$

$$\bar{a}_{10t}^s(\boldsymbol{a}_t, \varepsilon) = \chi\left(\varepsilon_{10t}^*, \varepsilon\right) \frac{1}{p_t} \left(a_t^m + p_t a_t^s\right). \tag{22}$$

(b) If the investor is able to contact only a bond broker, the post-trade portfolio is

$$\bar{a}_{01t}^m\left(\boldsymbol{a}_t\right) = \chi\left(1, q_t \phi_t^m\right) \left(a_t^m + \lambda q_t \phi_t^s a_t^s\right) \tag{23}$$

$$\bar{a}_{01t}^s\left(\boldsymbol{a}_t\right) = a_t^s \tag{24}$$

$$\bar{a}_{01t}^b\left(\boldsymbol{a}_t\right) = \left[1 - \chi\left(1, q_t \phi_t^m\right)\right] \frac{1}{q_t} a_t^m - \chi\left(1, q_t \phi_t^m\right) \lambda \phi_t^s a_t^s \tag{25}$$

and the intermediation fee is

$$k_{01t}\left(\boldsymbol{a}_t\right) = (1-\theta)\left(1 - q_t \phi_t^m\right) \left(\mathbb{I}_{\{q_t\phi_t^m<1\}} \frac{1}{q_t} a_t^m - \mathbb{I}_{\{1<q_t\phi_t^m\}} \lambda \phi_t^s a_t^s\right). \tag{26}$$

(c) If the investor is able to contact both an equity broker and a bond broker, the bargaining problem has a solution only if

$$\lambda < \frac{p_t}{q_t \phi_t^s}, \tag{27}$$

and in that case the post-trade portfolio is

$$\begin{aligned} \bar{a}_{11t}^m\left(\boldsymbol{a}_t, \varepsilon\right) = & \left\{\mathbb{I}_{\{1<q_t\phi_t^m\}} \left[1 - \chi\left(\varepsilon_{11t}^*, \varepsilon\right)\right] + \mathbb{I}_{\{q_t\phi_t^m=1\}} \mathbb{I}_{\left\{\varepsilon<\varepsilon_{11t}^*\right\}} \left[1 - \chi\left(q_t \phi_t^m, 1\right)\right]\right\} \left(a_t^m + p_t a_t^s\right) \\ & + \mathbb{I}_{\{q_t\phi_t^m=1\}} \mathbb{I}_{\left\{\varepsilon=\varepsilon_{11t}^*\right\}} \hat{a}_t^m \end{aligned} \tag{28}$$

$$\begin{aligned} \bar{a}_{11t}^s\left(\boldsymbol{a}_t, \varepsilon\right) = & \left\{\mathbb{I}_{\{q_t\phi_t^m=1\}} \mathbb{I}_{\left\{\varepsilon_{11t}^*<\varepsilon\right\}} + \left[1 - \mathbb{I}_{\{q_t\phi_t^m=1\}}\right] \chi\left(\varepsilon_{11t}^*, \varepsilon\right)\right\} \frac{a_t^m + p_t a_t^s}{p_t - \lambda q_t \phi_t^s} \\ & + \mathbb{I}_{\{q_t\phi_t^m=1\}} \mathbb{I}_{\left\{\varepsilon=\varepsilon_{11t}^*\right\}} \hat{a}_t^s \end{aligned} \tag{29}$$

$$\bar{a}_{11t}^b(\boldsymbol{a}_t, \varepsilon) = -\frac{1}{q_t} \left\{\left[\bar{a}_{11t}^m(\boldsymbol{a}_t, \varepsilon) - a_t^m\right] + p_t \left[\bar{a}_{11t}^s(\boldsymbol{a}_t, \varepsilon) - a_t^s\right]\right\}, \tag{30}$$

where

$$(\hat{a}_t^m, \hat{a}_t^s) \in \left\{ \mathbb{R}_+^2 : \hat{a}_t^m + (p_t - q_t \lambda \phi_t^s)\, \hat{a}_t^s \leq a_t^m + p_t a_t^s \right\},$$

and the intermediation fee is

$$k_{11t}(\boldsymbol{a}_t, \varepsilon) = (1-\theta) \Big\{ (\varepsilon y_t + \phi_t^s) \left[\overline{a}_{11t}^s(\boldsymbol{a}_t, \varepsilon) - \overline{a}_{10t}^s(\boldsymbol{a}_t, \varepsilon) \right] + \phi_t^m \left[\overline{a}_{11t}^m(\boldsymbol{a}_t, \varepsilon) - \overline{a}_{10t}^m(\boldsymbol{a}_t, \varepsilon) \right] + \overline{a}_{11t}^b(\boldsymbol{a}_t, \varepsilon) \Big\}. \tag{31}$$

(ii) The post-trade portfolio of a bond broker who enters the OTC round of period t with portfolio $\boldsymbol{a}_t$ is $\overline{\boldsymbol{a}}_{Bt}(\boldsymbol{a}_t) = \overline{\boldsymbol{a}}_{01t}(\boldsymbol{a}_t)$.

(iii) The post-trade portfolio of an equity broker who enters the OTC round of period t with portfolio $\boldsymbol{a}_t$ is $\overline{\boldsymbol{a}}_{Et}(\boldsymbol{a}_t) = \overline{\boldsymbol{a}}_{10t}(\boldsymbol{a}_t, 0)$.

Part $(i)(a)$ of Lemma 2 states that the individual investor with valuation ε who is only able to contact an equity broker in the OTC, uses all his money balances to buy equity if $\varepsilon_{10t}^* < \varepsilon$, and sells all his equity for money if $\varepsilon < \varepsilon_{10t}^*$. The intermediation fee is zero in this case given our assumption that the investor has all the bargaining power in negotiations with an equity broker. Part $(i)(b)$ considers the case of an investor with beginning-of-period asset holding $\boldsymbol{a}_t$ and valuation ε who only contacts a bond broker. Since this investor cannot trade equity, his post-trade equity holding is the same as the pre-trade equity holding. This investor may, however, buy or sell bonds for money. if $q_t \phi_t^m < 1$, then the investor uses all his money to buy bonds. Conversely, if $1 < q_t \phi_t^m$, the investor short sells as much of the bond as allowed by the collateral constraint, and holds the proceeds in the form of money. The bond broker is able to extract a fee whenever the investor's gain from trading bonds is positive. Part $(i)(c)$ describes the bargaining outcome of an investor who simultaneously contacts an equity and a bond broker. To offer a simple interpretation of the post-trade allocation in this case, suppose $q_t \phi_t^m < 1$ (which will in fact be the case in an equilibrium where the nominal interest rate on the bond is positive). Then if $\varepsilon_{11t}^* < \varepsilon$, the investor short sells as much of the bond as allowed by the collateral constraint, and uses the proceeds from the short sale along with all his pre-trade money balances to buy equity. Conversely, if $\varepsilon < \varepsilon_{11t}^*$, the investor sells all his pre-trade equity holding and uses the proceeds from the sale, and all his pre-trade money balances to buy bonds. The bond broker extracts a fee whenever the investor has positive gain from trade. Part (ii) states that the post-trade portfolio of a bond broker with some pre-trade portfolio is the same as the post-trade portfolio of an investor with the same pre-trade portfolio who only contacts a bond broker. Part (*iii*) states that the post-trade portfolio of an equity broker with some pre-trade portfolio is the same as the post-trade portfolio of an investor with the same pre-trade portfolio and valuation $\varepsilon = 0$ who only contacts an equity broker.

To see why condition (27) is necessary for the bargaining outcome of an investor of type 11 to be well defined, it is useful to think of the interest rate implied by the inside bond. First, notice that with 1 unit of money an investor can buy $\frac{1}{q_t}$ bonds, which in total yield $\frac{1}{q_t}$ general goods in the following subperiod, and this is equivalent to $\frac{1}{q_t\phi_t^m}$ dollars. Thus the gross nominal rate on a collateralized loan is

$$i_t^m \equiv \frac{1}{q_t\phi_t^m} - 1. \tag{32}$$

Since the loan is repaid within the period, this is also a notion of real rate on these loans, with loan and repayment measured in terms of the general good.[15] Another notion of interest rate that corresponds to the one we defined for the economy with no money, is the gross real interest rate on bonds expressed in terms of equity shares. If an investor uses 1 unit of money to purchase bonds, this is equivalent to investing $\frac{1}{p_t}$ worth of equity shares for $\frac{1}{q_t}$ bonds that deliver $\frac{1}{q_t}$ general goods in the following subperiod, and this is equivalent to $\frac{1}{q_t\phi_t^s}$ equity shares. So the real rate on a bond (in terms of equity shares) is

$$i_t^s \equiv \frac{p_t}{q_t\phi_t^s} - 1. \tag{33}$$

Consider the budget constraint and the collateral constraint of an investor who contacts an equity and a bond broker, namely $\overline{a}_t^m + p_t\overline{a}_t^s + q_t\overline{a}_t^b = a_t^m + p_t a_t^s$ and $-\lambda\phi_t^s\overline{a}_t^s \leq \overline{a}_t^b$. These two conditions imply the borrowing constraint $-\lambda\phi_t^s\left[a_t^s + \left(a_t^m - \overline{a}_t^m\right)/p_t\right] \leq \left(1 - \lambda q_t\phi_t^s/p_t\right)\overline{a}_t^b$. This constraint would be slack for all $\overline{a}_t^b < 0$ if (27) were violated, meaning that an investor with $\varepsilon > \varepsilon_{11t}^*$ would be able (and willing) to take an infinitely long position in the stock. Intuitively, notice that if (27) is violated, then an investor of type 11 who starts with no wealth can sell b collateralized bonds for money and use the monetary proceeds to purchase bq_t/p_t equity shares; this leveraged purchase would leave the investor's borrowing constraint slack, since $b < \lambda\phi_t^s bq_t/p_t$. Notice that with (33), (27) can be written as $\lambda < 1 + i_t^s$, so (27) could only be violated if the net real interest were negative, which as we show below, will not be the case in equilibrium.

[15] Investing $\frac{1}{\phi_t^m}$ dollars is equivalent to investing 1 unit of the general good. The $\frac{1}{\phi_t^m}$ dollars allow to buy $\frac{1}{q_t\phi_t^m}$ bonds, which in total yield $\frac{1}{q_t\phi_t^m}$ general goods. So the gross real interest in terms of general goods is also $\frac{1}{q_t\phi_t^m}$.

Hereafter, to simplify the exposition, unless otherwise specified, we assume that $\alpha_{01} = 0$ and $\alpha_{10}, \alpha_{11} \in (0,1)$, that brokers cannot hold equity or money overnight.[16] We focus the analysis on equilibria with an active bond market, which implies $q_t \phi_t^m \leq 1$ or equivalently, that the net nominal interest rate on bonds, i_t^m, is nonnegative.[17] In this case, (20) implies $\varepsilon_{11t}^* y_t = (p_t/q_t - \phi_t^s)$ and notice that $(\varepsilon_{11t}^* - \varepsilon_{10t}^*) y_t = (1/q_t - \phi_t^m) p_t \geq 0$. We consider recursive (stationary) equilibria, i.e., equilibria in which: (a) real equity prices (and real money balances, if the equilibrium is monetary) are time-invariant linear functions of the aggregate dividend, and (b) aggregate equity holdings of each agent type are constant over time. Next, we define nonmonetary, and monetary recursive equilibria in turn.

Definition 2 *A recursive nonmonetary equilibrium (RNE) is a nonmonetary equilibrium in which real equity prices (general goods per equity share) are time-invariant linear functions of the aggregate dividend, i.e., $\phi_t^s = \phi^s y_t$ and $\bar{\phi}_t^s = \bar{\phi}^s y_t$ for some $\phi^s, \bar{\phi}^s \in \mathbb{R}_+$, and aggregate equity holdings of each agent type are constant over time, i.e., $A_{jt}^s = A_j^s$ for $j \in \{B, E, I\}$.*

Hence in a RNE, we have $\varepsilon_t^n = \left(\bar{\phi}_t^s - \phi_t^s\right) \frac{1}{y_t} = \bar{\phi}^s - \phi^s \equiv \varepsilon^n$, and the real interest rate on the bond, i.e., $i_t^n \equiv \bar{\phi}_t^s / \phi_t^s - 1$ as defined in (18), is

$$i^n = \frac{\varepsilon^n}{\phi^s}. \tag{34}$$

Definition 3 *A recursive monetary equilibrium (RME) is a monetary equilibrium in which: (i) real equity prices (general goods per equity share) are time-invariant linear functions of the aggregate dividend, i.e., $\phi_t^s = \phi^s y_t$, $p_t \phi_t^m \equiv \bar{\phi}_{10t}^s = \bar{\phi}_{10}^s y_t$, and $p_t / q_t \equiv \bar{\phi}_{11t}^s = \bar{\phi}_{11}^s y_t$ for some $\phi^s, \bar{\phi}_{10}^s, \bar{\phi}_{11}^s \in \mathbb{R}_+$; (ii) aggregate equity holdings of each agent type are constant over time, i.e., $A_{jt}^s = A_j^s$ for $j \in \{B, E, I\}$; and (iii) real money balances are a constant proportion of output, i.e., $\phi_t^m A_t^m = Z A^s y_t$ for some $Z \in \mathbb{R}_{++}$, and $\phi_t^m A_{jt}^m = Z_j A^s y_t$, for $j \in \{B, E, I\}$ and $Z_j \in \mathbb{R}_+$.*

[16] We assume $\alpha_{01} = 0$ for two reasons. First, the analysis is cleanest in this case. Second, for the purposes of studying margin loans, it is sufficient and it is natural to assume that only investors who participate in equity trades get in contact with the brokers who offer margin loans. The assumption that brokers do not hold money overnight is immaterial (they would not want to hold it even if they were allowed), as is the assumption that bond brokers do not carry equity. Whether the assumption that equity brokers do not hold equity overnight represents a restriction on behavior will depend on the inflation rate; for inflation rates higher than a certain threshold, they would be disinclined to do so. See Lagos and Zhang (2015, 2017) for more details on these considerations in related trading environments.

[17] Lemma 10 in the appendix establishes that $q_t \phi_t^m \leq 1$ is necessary for bonds to trade in equilibrium. We focus on equilibria with an active bond market because if the bond is not traded, money is the only means of payment and the equilibrium conditions specialize to those in Lagos and Zhang (2016).

Hence in a RME, $\varepsilon_{10t}^{*} = (p_t\phi_t^m - \phi_t^s)\frac{1}{y_t} = \bar{\phi}_{10}^s - \phi^s \equiv \varepsilon_{10}^{*}$, $\varepsilon_{11t}^{*} = (p_t/q_t - \phi_t^s)\frac{1}{y_t} = \bar{\phi}_{11}^s - \phi^s \equiv \varepsilon_{11}^{*}$, $p_t = \frac{(\varepsilon_{10}^{*}+\phi^s)A_t^m}{ZA^s}$, $\phi_t^m = \frac{ZA^s y_t}{A_t^m}$, and

$$q_t = \frac{(\varepsilon_{10}^{*} + \phi^s)\, A_t^m}{(\varepsilon_{11}^{*} + \phi^s)\, ZA^s y_t}. \tag{35}$$

Thus, $\phi_{t+1}^s/\phi_t^s = \bar{\phi}_{10t+1}^s/\bar{\phi}_{10t}^s = \bar{\phi}_{11t+1}^s/\bar{\phi}_{11t}^s = \gamma_{t+1}$, $p_{t+1}/p_t = \mu$, and $\phi_t^m/\phi_{t+1}^m = q_{t+1}/q_t = \mu/\gamma_{t+1}$. In a RME, the (gross) interest rate on a collateralized loan (i_t^m as defined in (32)) is

$$i^m = \frac{\varepsilon_{11}^{*} - \varepsilon_{10}^{*}}{\varepsilon_{10}^{*} + \phi^s} \tag{36}$$

and the gross real (in terms of equity shares) interest rate on a collateralized loan (i_t^s as defined in (33)) is

$$i^s = \frac{\varepsilon_{11}^{*}}{\phi^s}. \tag{37}$$

Notice that in a RME, (27) becomes $\lambda < 1 + i^s$, which will be satisfied since $\lambda \in [0, 1]$.

Let $q_{t,k}^B$ denote the nominal price in the second subperiod of period t of an N-period risk-free pure discount nominal bond that matures in period $t+k$, for $k = 0, 1, 2, ..., N$ (so k is the number of periods until the bond matures). Imagine the bond is illiquid in the sense that it cannot be traded in the OTC market. Then in a stationary monetary equilibrium, $q_{t,k}^B = (\bar{\beta}/\mu)^k$, and

$$i^p = \frac{\mu - \bar{\beta}}{\bar{\beta}} \tag{38}$$

is the time t nominal yield to maturity of the bond with k periods until maturity. Throughout the analysis we let $\bar{\beta} \equiv \beta\bar{\gamma}$ and maintain the assumption $\mu > \bar{\beta}$ (but we consider the limiting case $\mu \to \bar{\beta}$). Since there is a one-to-one mapping between the policy variable μ and the interest rate i^p, we can regard i^p as the nominal *policy rate* chosen by the monetary authority.

Hereafter, we focus on a formulation of the model where the length of the time period becomes arbitrarily short, which allows us to deliver sharp theoretical results. This limiting economy can be interpreted as an approximation to a continuous-time version of our discrete-time economy. To this end, we first generalize the discrete-time model by allowing the period length to be an arbitrary constant, and then take the limit as this constant becomes arbitrarily small. Let Δ denote the length of the model period, and define the discount rate, r, the expected dividend growth rate, g, the depreciation rate, δ, and the money growth rate, π, as $\beta \equiv (1 + r\Delta)^{-1}$, $\bar{\gamma} \equiv 1 + g\Delta$, $\eta \equiv 1 - \delta\Delta$, $\mu \equiv 1 + \pi\Delta$. Over a time period of length Δ, the dividend is $y_t\Delta$, and consumption of the dividend good is $\varepsilon y_t\Delta$. In this context we focus on recursive equilibria where, as $\Delta \to 0$, real asset prices are time-invariant linear functions of the *dividend rate*, y_t. Specifically, let $\Phi_t^s(\Delta)$ and $\Phi_t^m(\Delta) A_t^m$ denote the real equity price and the real aggregate money balance, respectively, in the discrete-time economy with time periods of length Δ. We look for recursive equilibria of this discrete-time economy such that $\Phi_t^s(\Delta) = \Phi^s(\Delta) y_t \Delta$ and $\Phi_t^m(\Delta) A_t^m = \mathrm{Z}(\Delta) A^s y_t \Delta$, where $\Phi^s(\Delta)$ and $\mathrm{Z}(\Delta)$ are time-invariant functions with the property that $\lim_{\Delta \to 0} \Phi^s(\Delta)\Delta = \phi^s$ and $\lim_{\Delta \to 0} \mathrm{Z}(\Delta)\Delta = Z$, with $\phi^s, Z \in \mathbb{R}$. Hence (34), (36), (37), and (38) generalize to $i^n = \frac{\varepsilon^n}{\Phi^s(\Delta)}$, $i^m = \frac{\varepsilon_{11}^* - \varepsilon_{10}^*}{\varepsilon_{10}^* + \Phi^s(\Delta)}$, $i^s = \frac{\varepsilon_{11}^*}{\Phi^s(\Delta)}$, and $i^p = \frac{(r+\pi-g+r\pi\Delta)\Delta}{1+g\Delta}$, respectively.[18]

Define

$$\rho^n \equiv \lim_{\Delta \mapsto 0} \frac{i^n}{\Delta} = \frac{\varepsilon^n}{\phi^s}$$
$$\rho^m \equiv \lim_{\Delta \mapsto 0} \frac{i^m}{\Delta} = \frac{\varepsilon_{11}^* - \varepsilon_{10}^*}{\phi^s}$$
$$\rho^s \equiv \lim_{\Delta \mapsto 0} \frac{i^s}{\Delta} = \frac{\varepsilon_{11}^*}{\phi^s}$$
$$\rho^p \equiv \lim_{\Delta \mapsto 0} \frac{i^p}{\Delta} = r + \pi - g.$$

Intuitively, ρ^n is the real interest rate on inside bonds in a nonmonetary economy, ρ^m and ρ^s are the nominal and real interest rates on inside bonds in a monetary economy, and ρ^p is the policy nominal interest rate that can be controled by the monetary authority (e.g., by changing π). Notice that $\rho^p = r + \bar{\pi}$ is a Fisher equation that equates the nominal rate interest rate (on an illiquid outside bond) to the real risk-free interest rate, r, plus an expected inflation rate, $\bar{\pi} \equiv \pi - g$ (measured with the price of the general consumption good).[19] For what follows, it is useful to let $\varphi \equiv \rho\phi^s$, $\mathcal{Z} \equiv \rho Z$, and $\iota \equiv \rho^p/\rho$, where

$$\rho \equiv \lim_{\Delta \mapsto 0} \frac{1}{\Delta} \frac{1 - \bar{\beta}\eta}{\bar{\beta}\eta} = r + \delta - g.$$

The factor ρ can be interpreted as a *capitalization rate* for equity holdings, so φ and $\mathcal{Z}$ are the "flow values" of an equity share and a unit of real balances, respectively, and ι is the policy

[18] The discrete-time formulation we laid out previously, corresponds to a special case of this formulation with $\Delta = 1$, $\Phi_t^s(1) \equiv \phi_t^s$, $\Phi_t^m(1) \equiv \phi_t^m$, $\Phi^s(1) \equiv \phi^s$, and $\mathrm{Z}(1) \equiv Z$.

[19] The (gross) inflation rate in terms of general goods is $\phi_t^m/\phi_{t+1}^m = \mu \frac{y_t}{y_{t+1}} \equiv 1 + \tilde{\pi}_{t+1}$, which is stochastic. A measure of this average (expected) inflation is

$$\left[\mathbb{E}_t \frac{1}{1+\tilde{\pi}_{t+1}}\right]^{-1} = \frac{\mu}{\bar{\gamma}} = \frac{1+\pi\Delta}{1+g\Delta} \equiv 1 + \bar{\pi}\Delta,$$

so as $\Delta \to 0$, we get $\bar{\pi} = \pi - g$ as a measure of the expected inflation in the nominal price of general goods. Since $p_{t+1}/p_t - 1 = \mu - 1 = \pi\Delta$, the inflation rate in the nominal price of equity shares is π.

rate up to a convenient normalization. In the remainder, we focus on the limiting economy that obtains as $\Delta \to 0$.

Proposition 1 *Consider the limiting economy (as $\Delta \to 0$). There exists a unique recursive nonmonetary equilibrium, $(\varepsilon^n, \varphi^n)$. Moreover,*

$$\varphi^n = \bar{\varepsilon} + \alpha_{11}\theta \left[\int_{\varepsilon_L}^{\varepsilon^n} (\varepsilon^n - \varepsilon)\, dG(\varepsilon) + \frac{\lambda}{1-\lambda} \int_{\varepsilon^n}^{\varepsilon_H} (\varepsilon - \varepsilon^n)\, dG(\varepsilon) \right] \tag{39}$$

and $\varepsilon^n \in [\varepsilon_L, \varepsilon_H]$ is the unique solution to

$$G(\varepsilon^n) = \lambda. \tag{40}$$

The asset price in the nonmonetary equilibrium can be decomposed into three components. The first term in (39) is the expected value of the dividend flow. The second term in (39) is the expected gain from exercising the option of reselling the asset in the OTC market. The third term in (39) reflects the expected marginal value of the asset when it is pledged as collateral for shorting the bond. We will label these components the *(value of the) resale option*, and the *(value of the) pledge option*, respectively. From (40), it is clear the marginal investor valuation that clears the OTC market, ε^n, is increasing in λ, reflecting the fact that as λ increases, the collateral constraint is relaxed, and higher valuation investors are able to absorb a higher proportion of the asset holdings. In particular, $\varepsilon^n \to \varepsilon_H$ as $\lambda \to 1$. For a given ε^n, the asset price is increasing in λ, but in the general equilibrium, a larger λ implies a larger ε^n, which in turn implies an investor is less likely to receive a valuation shock large enough to want to use the asset as collateral, and this force tends to make the price decreasing in λ. It is possible to show, however, that the former effect always dominates, and φ^n is increasing in λ even after taking into acccount the general equilibrium effect (see Lemma 14 in the appendix).

The model is quite general in terms of the trading situations investors may face in the OTC round: A fraction α_{10} of investors trade equity with no access to a margin loan, a fraction α_{11} trade equity with access to margin loans, and a fraction α_{01} do not participate of the equity market but are able to lend money through bond brokers. The following proposition offers a complete characterization of the set of RME for an economy with $\alpha_{01} = 0$. Let

$$\bar{\iota}(\lambda) \equiv \frac{\alpha_{11}\theta(\varepsilon^n - \varepsilon_L) + [\alpha_{10} + \alpha_{11}(1-\theta)](\bar{\varepsilon} - \varepsilon_L) + \alpha_{11}\theta\frac{1}{1-\lambda}\int_{\varepsilon^n}^{\varepsilon_H}(\varepsilon - \varepsilon^n)\, dG(\varepsilon)}{\bar{\varepsilon} + \alpha_{11}\theta\left[\int_{\varepsilon_L}^{\varepsilon^n}(\varepsilon^n - \varepsilon)\, dG(\varepsilon) + \frac{\lambda}{1-\lambda}\int_{\varepsilon^n}^{\varepsilon_H}(\varepsilon - \varepsilon^n)\, dG(\varepsilon)\right]} \tag{41}$$

$$\hat{\iota}(\lambda) \equiv \frac{\left[\alpha_{10} + \alpha_{11}\left(1 + \theta\frac{\lambda}{1-\lambda}\right)\right]\int_{\varepsilon^n}^{\varepsilon_H}(\varepsilon - \varepsilon^n)\, dG(\varepsilon)}{\bar{\varepsilon} + (\alpha_{10} + \alpha_{11})\int_{\varepsilon_L}^{\varepsilon^n}(\varepsilon^n - \varepsilon)\, dG(\varepsilon) + \alpha_{11}\theta\frac{\lambda}{1-\lambda}\int_{\varepsilon^n}^{\varepsilon_H}(\varepsilon - \varepsilon^n)\, dG(\varepsilon)}. \tag{42}$$

Proposition 2 *Consider the limiting economy (as $\Delta \to 0$).*

(i) If $\hat{\iota}(\lambda) < \iota < \bar{\iota}(\lambda)$ then there exists a unique recursive monetary equilibrium, $(\varepsilon_{10}^*, \varepsilon_{11}^*, \varphi, \mathcal{Z})$. The asset prices are

$$\varphi = \varphi^n + [\alpha_{10} + \alpha_{11}(1-\theta)]\int_{\varepsilon_L}^{\varepsilon_{10}^*}(\varepsilon_{10}^* - \varepsilon)\, dG(\varepsilon) \tag{43}$$

$$\mathcal{Z} = \frac{\alpha_{10}G(\varepsilon_{10}^*)}{[1 - G(\varepsilon_{10}^*)]\,\alpha_{10} + \alpha_{11}}\varphi. \tag{44}$$

The marginal valuations are $\varepsilon_{11}^* = \varepsilon^n$ and the unique $\varepsilon_{10}^* \in (\varepsilon_L, \varepsilon^n)$ that satisfies

$$\frac{\alpha_{11}\theta\left(\varepsilon^n - \varepsilon_{10}^*\right) + [\alpha_{10} + \alpha_{11}(1-\theta)]\int_{\varepsilon_{10}^*}^{\varepsilon_H}\left(\varepsilon - \varepsilon_{10}^*\right) dG(\varepsilon) + \alpha_{11}\theta\frac{1}{1-\lambda}\int_{\varepsilon^n}^{\varepsilon_H}(\varepsilon - \varepsilon^n) dG(\varepsilon)}{\varphi^n + [\alpha_{10} + \alpha_{11}(1-\theta)]\int_{\varepsilon_L}^{\varepsilon_{10}^*}\left(\varepsilon_{10}^* - \varepsilon\right) dG(\varepsilon)} = \iota.$$

(ii) If $0 < \iota \leq \hat{\iota}(\lambda)$ then there exists a unique recursive monetary equilibrium, $(\varepsilon^*, \chi, \varphi, \mathcal{Z})$. The asset prices are

$$\varphi = \bar{\varepsilon} + (\alpha_{10} + \alpha_{11})\int_{\varepsilon_L}^{\varepsilon^*}(\varepsilon^* - \varepsilon)\, dG(\varepsilon) + \alpha_{11}\theta\frac{\lambda}{1-\lambda}\int_{\varepsilon^*}^{\varepsilon_H}(\varepsilon - \varepsilon^*)\, dG(\varepsilon)$$

$$\mathcal{Z} = \frac{\alpha_{10}G(\varepsilon^*) + \alpha_{11}\frac{1}{1-\lambda}[G(\varepsilon^*) - \lambda]}{[1 - G(\varepsilon^*)]\left(\alpha_{10} + \alpha_{11}\frac{1}{1-\lambda}\right)}\varphi.$$

The marginal valuations are $\varepsilon_{10}^* = \varepsilon_{11}^* \equiv \varepsilon^*$, where $\varepsilon^* \in [\varepsilon^n, \varepsilon_H)$ (with $\varepsilon^* = \varepsilon^n$ only if $\iota = \hat{\iota}(\lambda)$) is the unique solution to

$$\frac{\left[\alpha_{10} + \alpha_{11}\left(1 + \theta\frac{\lambda}{1-\lambda}\right)\right]\int_{\varepsilon^*}^{\varepsilon_H}(\varepsilon - \varepsilon^*)\, dG(\varepsilon)}{\bar{\varepsilon} + (\alpha_{10} + \alpha_{11})\int_{\varepsilon_L}^{\varepsilon^*}(\varepsilon^* - \varepsilon)\, dG(\varepsilon) + \alpha_{11}\theta\frac{\lambda}{1-\lambda}\int_{\varepsilon^*}^{\varepsilon_H}(\varepsilon - \varepsilon^*)\, dG(\varepsilon)} = \iota$$

and

$$\chi = \frac{\lambda}{1-\lambda}\frac{1 - G(\varepsilon^*)}{G(\varepsilon^*)}$$

is the proportion of the financial wealth that investors of type 11 with valuation lower than ε^* hold in the form of bonds (they hold the remaining $1 - \chi$ fraction in cash).

Figure 1 illustrates the existence regions in the space of parameters ι (vertical axis) and λ (horizontal axis). The boundaries $\iota = \bar{\iota}(\lambda)$ and $\iota = \hat{\iota}(\lambda)$ define three regions.[20] As shown in Proposition 1, a nonmonetary equilibrium exists for any parametrization, and therefore for every configuration of parameters shown in Figure 1. For each $\lambda \in [0,1]$, $\bar{\iota}(\lambda)$ is the largest nominal policy rate consistent with existence of monetary equilibrium. No monetary equilibrium exists for $\iota \geq \bar{\iota}(\lambda)$, so in that case the nonmonetary equilibrium is the unique equilibrium. For every $\iota < \bar{\iota}(\lambda)$, the nonmonetary equilibrium coexists with a monetary equilibrium. For each $\lambda \in (0,1]$, the monetary equilibrium is qualitatively different depending on whether the policy rate is higher or lower than $\hat{\iota}(\lambda)$.

If $\hat{\iota}(\lambda) < \iota < \bar{\iota}(\lambda)$, then credit is scarce in the sense that the nominal interest rate on margin loans, ρ^m, is positive. In this case, type 10 investors with valuations larger than ε_{10}^* use all their money to take a long position in equity, and investors of type 11 with valuations larger than ε_{11}^* short the bond up to the collateral constraint, and use all their money in order to purchase equity. Investors of type 10 with valuations smaller than ε_{10}^* sell all their equity holdings for money, and investors of type 11 with valuations smaller than ε_{11}^* sell all their equity holdings and use all the proceeds along with their pre-trade money holdings to take a long position in bonds. Notice that in this parameter range, $\varepsilon_L < \varepsilon_{10}^* < \varepsilon_{11}^* = \varepsilon^n$, so the marginal equity buyer of type 10 has a lower valuation than the marginal equity buyer of type 11, and the latter has the same valuation he would have in the nonmonetary equilibrium. The fact that $\varepsilon_{10}^* < \varepsilon_{11}^*$ reflects that type 11 investors have the option of investing money in the interest-yielding inside bond, while this option is not available to investors of type 10. Notice that the equity price in this region of the parameter space is larger than the equity price in the nonmonetary equilibrium. The first reason (captured by the term multiplied by α_{10} in (43)) is that in a monetary equilibrium there are equity-for-money trades in the equity market so the investor of type 10 can resell equity if her valuation is relatively low, while this is not possible in the nonmonetary equilibrium. The second reason (captured by the term multiplied by $\alpha_{11}(1-\theta)$ in (43)) is that the option of being able to sell equity for money in the equity market improves the outside option of a type

[20] It is easy to prove that $\hat{\iota}(\lambda) \leq \bar{\iota}(\lambda)$ for all $\lambda \in [0,1]$ (with "=" only if $\lambda = 0$), and that $\hat{\iota}(0) = \bar{\iota}(0) = \frac{\alpha_{10}+\alpha_{11}}{\bar{\varepsilon}}(\bar{\varepsilon} - \varepsilon_L)$, $\hat{\iota}(1) = 0$, and

$$\bar{\iota}(1) = \frac{\alpha_{11}\theta(\varepsilon_H - \varepsilon_L) + [\alpha_{10} + \alpha_{11}(1-\theta)](\bar{\varepsilon} - \varepsilon_L)}{\bar{\varepsilon} + \alpha_{11}\theta(\varepsilon_H - \bar{\varepsilon})}.$$

Hence, $\hat{\iota}(1) < \bar{\iota}(0) = \hat{\iota}(0) \leq \bar{\iota}(1)$. Also, $\bar{\iota}(1) - \bar{\iota}(0) = \alpha_{11}\theta(\varepsilon_H - \bar{\varepsilon})\frac{(\alpha_{10}+\alpha_{11})\varepsilon_L + (1-\alpha_{10}-\alpha_{11})\bar{\varepsilon}}{[\bar{\varepsilon}+\alpha_{11}\theta(\varepsilon_H - \bar{\varepsilon})]\bar{\varepsilon}}$, so $0 \leq \bar{\iota}(1) - \bar{\iota}(0)$ holds with "=" only if $\alpha_{11}\theta = 0$. In particular, notice $\theta = 0$ (or $\alpha_{11} = 0$) implies $\bar{\iota}(\lambda) = \bar{\iota}(0)$ for all λ.

11 investor when he bargains for the terms of the trade with the bond broker who helps the investor take a long position in the bond.

Conversely, if $0 < \iota \leq \hat{\iota}(\lambda)$, then real balances are abundant and credit demand is relatively weak, so $\rho^m = 0$ and money is not dominated by inside bonds in terms of rate of return. In this case, investors of type 10 with valuations larger than ε^* use all their money to buy equity while investors of type 10 with valuations lower than ε^* sell all their equity for money. Investors of type 11 with valuations larger than ε^* short the bond up to the collateral constraint, and use all their money in order to purchase equity. Investors of type 11 with valuations smaller than ε^*_{11} sell all their equity holdings, and are indifferent between holding the proceeds (and any pre-trade money holdings) in the form of money or bonds. Notice that in this parameter range, $\varepsilon_L < \varepsilon^n \leq \varepsilon^*_{10} = \varepsilon^*_{11} = \varepsilon^* < \varepsilon_H$, so the marginal equity buyer of type 10 has the same valuation as the marginal equity buyer of type 11, and this valuation is larger than the valuation an investor of type 11 would have in the nonmonetary equilibrium. The fact that $\varepsilon^*_{10} = \varepsilon^*_{11}$ reflects that the nominal rate on inside bonds is equal to zero, so type 11 investors and type 10 investors have the same valuation for money in the OTC round, even though only the former can lend it in the bond market.

Interestingly, the monetary equilibrium becomes a more robust trading arrangement when the loan-to-value ratio λ increases. To see this, notice that

$$\frac{\partial \bar{\iota}(\lambda)}{\partial \lambda} = \frac{\alpha_{11}\theta\left[\bar{\varepsilon} - (\alpha_{10} + \alpha_{11})(\bar{\varepsilon} - \varepsilon_L)\right]\int_{\varepsilon^n}^{\varepsilon_H}(\varepsilon - \varepsilon^n)\,dG(\varepsilon)}{\left\{\left[1 - \lambda(1 - \alpha_{11}\theta)\right]\bar{\varepsilon} + \alpha_{11}\theta\left[\int_{\varepsilon_L}^{\varepsilon^n}(\varepsilon^n - \varepsilon)\,dG(\varepsilon) - \lambda\hat{\varepsilon}\right]\right\}^2} \geq 0$$

with "=" only if $\alpha_{11}\theta = 0$. In other words, the maximum rate of inflation consistent with monetary equilibrium is increasing with λ. So as long as $0 < \alpha_{11}\theta$, money and credit (in the form of leveraged purchases of equity) behave as complements, rather than substitutes, in the general equilibrium. The reason money has value in the monetary equilibrium is that high-valuation investors use it to finance a long postition in the equity. So for larger λ, this value of money in exchange is enhanced because the equity shares an investor of type 11 buys with money can themselves be used as collateral to short the bond and take an even larger long position in the equity.

The following result summarizes the implications of monetary exchange and monetary policy for asset prices.

Proposition 3 *(i) The real asset price in the monetary equilibrium is higher than in the non-monetary equilibrium, i.e.,* $\varphi^n \leq \varphi$ *for all* $\iota \in [0, \bar{\iota}(\lambda)]$*, with "=" only if* $\iota = \bar{\iota}(\lambda)$*. Moreover,* $\varphi \leq \psi \equiv \bar{\varepsilon} + (\alpha_{10} + \alpha_{11})(\varepsilon_H - \bar{\varepsilon})$*, with "=" only if* $\iota = 0$*.*

(ii) In a monetary equilibrium, φ *is decreasing in* ι*. If* $\hat{\iota}(\lambda) < \iota < \bar{\iota}(\lambda)$*, then* ε^*_{10} *(the valuation of the marginal investor with no access to credit) is decreasing in* ι*. If* $0 < \iota \leq \hat{\iota}(\lambda)$*, then* ε^* *(the valuation of the marginal investor) is decreasing in* ι*.*

3.1 Example

In this section we consider the special case $G(\varepsilon) = G_L \mathbb{I}_{\{\varepsilon_L \leq \varepsilon < \varepsilon_H\}} + \mathbb{I}_{\{\varepsilon_H \leq \varepsilon\}}$ with $G_L = 1 - G_H \in [0, 1]$, for which equilibrium can be solved for in closed form. The explicit formulas for equilibrium allocations and prices allow us to clearly explain the key mechanisms. The following proposition summarizes the characterization for the RNE.

Proposition 4 *Assume* $G(\varepsilon) = G_L \mathbb{I}_{\{\varepsilon_L \leq \varepsilon < \varepsilon_H\}} + \mathbb{I}_{\{\varepsilon_H \leq \varepsilon\}}$*, and consider the limiting economy (as* $\Delta \to 0$*) with* $\alpha_{10}, \alpha_{11}, \alpha_{01} \in (0, 1)$*. There exists a unique stationary nonmonetary equilibrium (generically, for any combination of parameters with nonzero measure). The equity price is*

$$\varphi^n = \begin{cases} \bar{\varepsilon} + \alpha_{11}\theta\pi_H(\varepsilon_H - \varepsilon_L)\frac{\lambda}{1-\lambda} & \lambda < G_L \\ \bar{\varepsilon} + \alpha_{11}\theta\pi_L(\varepsilon_H - \varepsilon_L) & G_L \leq \lambda \end{cases}$$

and the valuation of the investor who is indifferent between purchasing equity or lending is

$$\varepsilon^n \begin{cases} = \varepsilon_L & \text{if } \lambda < G_L \\ \in [\varepsilon_L, \varepsilon_H] & \text{if } \lambda = G_L \\ = \varepsilon_H & \text{if } G_L < \lambda. \end{cases}$$

The equilibrium in Proposition 4 depends qualitatively on the degree to which investors can buy on margin. If $\lambda < G_L$ (margins are high) then the equity price consists of the dividend flow plus the value of the pledge option (the value of the resale option is nil because the marginal investor of type 11 who is pricing the asset has valuation ε_L). In this case the asset price is increasing in λ. If $G_L \leq \lambda$ then the equity price consists of the dividend flow plus the value of the resale option (the value of the pledge option is nil because the marginal investor of type 11 who is pricing the asset has valuation ε_H). In this case the asset price is independent of λ. The (net) real interest rate in the nonmonetary equilibrium is

$$\rho^n = \frac{\varepsilon^n}{\varphi^n}\rho \begin{cases} = \frac{\varepsilon_L}{\bar{\varepsilon}+\alpha_{11}\theta G_H(\varepsilon_H-\varepsilon_L)\frac{\lambda}{1-\lambda}}\rho & \text{if } \lambda < G_L \\ \in \left[\frac{\varepsilon_L}{\bar{\varepsilon}+\alpha_{11}\theta\pi_L(\varepsilon_H-\varepsilon_L)}\rho, \frac{\varepsilon_H}{\bar{\varepsilon}+\alpha_{11}\theta\pi_L(\varepsilon_H-\varepsilon_L)}\rho\right] & \text{if } \lambda = G_L \\ = \frac{\varepsilon_H}{\bar{\varepsilon}+\alpha_{11}\theta G_L(\varepsilon_H-\varepsilon_L)}\rho & \text{if } G_L < \lambda. \end{cases}$$

Notice ρ^n is nonmonotonic; specifically, it is decreasing in λ for all $\lambda < G_L$, but

$$\lim_{\lambda \uparrow G_L} \rho^n = \frac{\varepsilon_L}{\bar{\varepsilon} + \alpha_{11}\theta G_L (\varepsilon_H - \varepsilon_L)} \rho < \frac{\varepsilon_H}{\bar{\varepsilon} + \alpha_{11}\theta G_L (\varepsilon_H - \varepsilon_L)} \rho = \lim_{\lambda \downarrow G_L} \rho^n.$$

Let $\bar{A}_{ij}^{sk}$ denote the total number of equity shares held by investors of type ij with valuation ε_k for $k \in \{L, H\}$, at the end of the OTC round of trade. From Lemma 1, the aggregate post-trade portfolios at the end of the OTC round are $\bar{A}_{11}^{sL} = \alpha_{11} G_L \frac{1}{1-\lambda} \chi(\varepsilon^n, \varepsilon_L) A^s$ and $\bar{A}_{11}^{sH} = \alpha_{11} G_H \frac{1}{1-\lambda} \chi(\varepsilon^n, \varepsilon_H) A^s$. Thus in the nonmonetary equilibrium, $\bar{A}_{10}^{sL} = \bar{A}_{10}^{sH} = \alpha_{10} A^s$, and

$$\bar{A}_{11}^{sL} = \mathbb{I}_{\{\varepsilon^n = \varepsilon_L\}} \alpha_{11} \frac{G_L - \lambda}{1 - \lambda} A^s$$
$$\bar{A}_{11}^{sH} = \begin{cases} \alpha_{11} \frac{G_H}{1-\lambda} A^s & \text{if } \varepsilon^n < \varepsilon_H \\ \alpha_{11} A^s & \text{if } \varepsilon^n = \varepsilon_H. \end{cases}$$

In order to describe monetary equilibria, it is useful to introduce some notation. Let

$$\bar{\iota} \equiv \frac{(\alpha_{01} + \alpha_{11} G_L)\theta + (\alpha_{10} + \alpha_{11}) G_H}{\bar{\varepsilon} + \alpha_{11}\theta G_L (\varepsilon_H - \varepsilon_L)} (\varepsilon_H - \varepsilon_L) \tag{45}$$

$$\hat{\iota}(\lambda) \equiv \frac{\left[\alpha_{10} + \alpha_{11}\left(1 + \theta \frac{\lambda}{1-\lambda}\right)\right] G_H (\varepsilon_H - \varepsilon_L)}{\bar{\varepsilon} + \alpha_{11}\theta G_H \frac{\lambda}{1-\lambda} (\varepsilon_H - \varepsilon_L)} \tag{46}$$

$$\hat{\lambda} \equiv \frac{(\alpha_{01} + \alpha_{11})(\alpha_{10} + \alpha_{11}) G_L}{(\alpha_{01} + \alpha_{11})(\alpha_{10} + \alpha_{11}) - \alpha_{10}\alpha_{01} G_H}. \tag{47}$$

Proposition 5 *Assume $G(\varepsilon) = G_L \mathbb{I}_{\{\varepsilon_L \le \varepsilon < \varepsilon_H\}} + \mathbb{I}_{\{\varepsilon_H \le \varepsilon\}}$. Consider the limiting economy (as $\Delta \to 0$) with $\alpha_{10}, \alpha_{11}, \alpha_{01} \in (0,1)$. Define the parameter regions*

$$\mathcal{E}_1^m = \left\{(\iota, \lambda) \in [0, \bar{\iota}] \times [0,1] : \iota < \bar{\iota} \text{ and } \hat{\lambda} < \lambda\right\}$$
$$\mathcal{E}_2^m = \left\{(\iota, \lambda) \in [0, \bar{\iota}] \times [0,1] : 0 < \iota < \hat{\iota}(\lambda) \text{ and } \lambda < \hat{\lambda}\right\}$$
$$\mathcal{E}_3^m = \left\{(\iota, \lambda) \in [0, \bar{\iota}] \times [0,1] : \hat{\iota}(\lambda) < \iota < \bar{\iota} \text{ and } G_L < \lambda < \hat{\lambda}\right\}.$$

There exists a unique stationary monetary equilibrium for any combination of parameters with nonzero measure such that (ι, λ) belongs to the closure of $\mathcal{E}_1^m \cup \mathcal{E}_2^m \cup \mathcal{E}_3^m$.

(i) If $(\iota,\lambda)\in\mathcal{E}_1^m$*, then* $\varepsilon_L<\varepsilon_{10}^*<\varepsilon_{11}^*=\varepsilon_H$*, with*

$$\varepsilon_{10}^* = \varepsilon_L + \frac{\bar{\iota}-\iota}{\frac{\bar{\iota}}{\varepsilon_H-\varepsilon_L}+\iota\frac{[\alpha_{10}+\alpha_{11}(1-\theta)]G_L}{\bar{\varepsilon}+\alpha_{11}\theta G_L(\varepsilon_H-\varepsilon_L)}} \tag{48}$$

$$\varphi = \frac{\bar{\varepsilon}+(\alpha_{10}+\alpha_{11})\,G_L\,(\varepsilon_H-\varepsilon_L)}{1+\iota\frac{[\alpha_{10}+\alpha_{11}(1-\theta)]G_L}{(\alpha_{01}+\alpha_{11})\theta+[\alpha_{10}+\alpha_{11}(1-\theta)]G_H}} \tag{49}$$

$$\mathcal{Z} = \frac{G_L\alpha_{10}}{\alpha_{01}+G_H\alpha_{10}+\alpha_{11}}\varphi. \tag{50}$$

(ii) If $(\iota,\lambda)\in\mathcal{E}_2^m$*, then* $\varepsilon_L<\varepsilon_{10}^*=\varepsilon_{11}^*\equiv\varepsilon^*<\varepsilon_H$*, with*

$$\varepsilon^* = \varepsilon_H - \frac{\iota\,[\bar{\varepsilon}+(\alpha_{10}+\alpha_{11})\,G_L\,(\varepsilon_H-\varepsilon_L)]}{\left[\alpha_{10}+\alpha_{11}\left(1+\theta\frac{\lambda}{1-\lambda}\right)\right]G_H+\iota\left\{\alpha_{10}G_L+\alpha_{11}\left[1-G_H\left(1+\theta\frac{\lambda}{1-\lambda}\right)\right]\right\}} \tag{51}$$

$$\varphi = \frac{\bar{\varepsilon}+(\alpha_{10}+\alpha_{11})\,G_L\,(\varepsilon_H-\varepsilon_L)}{1+\iota\frac{\alpha_{10}G_L+\alpha_{11}\left[1-G_H\left(1+\theta\frac{\lambda}{1-\lambda}\right)\right]}{\left[\alpha_{10}+\alpha_{11}\left(1+\theta\frac{\lambda}{1-\lambda}\right)\right]G_H}} \tag{52}$$

$$\mathcal{Z} = \left[\frac{\alpha_{10}+\alpha_{11}}{\left(\alpha_{10}+\alpha_{11}\frac{1}{1-\lambda}\right)G_H}-1\right]\varphi. \tag{53}$$

(iii) If $(\iota,\lambda)\in\mathcal{E}_3^m$*, then* $\varepsilon_L=\varepsilon_{10}^*<\varepsilon_{11}^*<\varepsilon_H$*, with*

$$\varepsilon_{11}^* = \varepsilon_H - \frac{(\bar{\iota}-\iota)\,[\bar{\varepsilon}+\alpha_{11}\theta G_L\,(\varepsilon_H-\varepsilon_L)]}{\theta\left[\alpha_{01}+(\iota-1)\,\alpha_{11}\left(G_H\frac{\lambda}{1-\lambda}-G_L\right)\right]} \tag{54}$$

$$\varphi = \frac{(1-\bar{\iota})\,\alpha_{11}\left(G_H\frac{\lambda}{1-\lambda}-G_L\right)-\alpha_{01}}{(1-\iota)\,\alpha_{11}\left(G_H\frac{\lambda}{1-\lambda}-G_L\right)-\alpha_{01}}\,[\bar{\varepsilon}+\alpha_{11}\theta G_L\,(\varepsilon_H-\varepsilon_L)] \tag{55}$$

$$\mathcal{Z} = \frac{\alpha_{11}\left(G_H\frac{\lambda}{1-\lambda}-G_L\right)}{\alpha_{01}-\alpha_{11}\left(G_H\frac{\lambda}{1-\lambda}-G_L\right)}\varphi. \tag{56}$$

According to Proposition 5, monetary equilibrium exists if and only if $(\iota,\lambda)\in\mathcal{E}_1^m\cup\mathcal{E}_2^m\cup\mathcal{E}_3^m$. Economies with $(\iota,\lambda)\in\mathcal{E}_1^m$ can be thought of as high-leverage economies. Economies with $(\iota,\lambda)\in\mathcal{E}_2^m$ have relatively low leverage and relatively low nominal policy rate. Economies with $(\iota,\lambda)\in\mathcal{E}_3^m$ have relatively low leverage and relatively high nominal policy rate (the set $\mathcal{E}_3^m$ collapses to a point if $\alpha_{01}=0$). We describe the monetary equilibrium of each of these economies in turn.

In an economy $\mathcal{E}_1^m$, investor's post trade portfolios in the OTC market are as follows: (a) since $\varepsilon_{10}^*<\varepsilon_H$, high-valuation investors of type 10 hold equity and no money (and no bonds, since they cannot access the bond market); (b) since $\varepsilon_L < \varepsilon^*_{10}$, low-valuation investors of type 10 hold money and no equity (and no bonds); (c) a fraction $\chi^{sH}_{11} = \frac{(1-\lambda)[\alpha_{11}(\alpha_{01}+\alpha_{10}+\alpha_{11})+\alpha_{10}\alpha_{01}G_L]}{\alpha_{11}(\alpha_{01}+\alpha_{10}+\alpha_{11})G_H}$ of the high-valuation investors of type 11 carry equity, a negative bond position, and no money; (d) the remaining $1-\chi^{sH}_{11}$ fraction of high-valuation investors of type 11, and low-valuation investors of type 11, carry a positive bond position, no money, and no equity; (e) investors of type 01 hold a positive bond position, no money, and the equity they had at the beginning of the period. Notice that all high-valuation investors of type 10 hold equity, while not all high-valuation investors of type 11 hold equity. The reason is that the equilibrium nominal interest rate on inside bonds, $\rho^m = \frac{\varepsilon_H - \varepsilon^*_{10}}{\varphi}\rho$ is positive and high enough to induce some type 11 high-valuation investors to buy bonds rather than equity (both yield the same return in equilibrium). In contrast, type 10 investors have no other use for their money than investing in equity.

In economy $\mathcal{E}^m_2$, investor's post trade portfolios in the OTC market are as follows: (a) high-valuation investors of type 10 carry equity, no money, and no bonds; (b) low-valuation investors of type 10 carry money, no equity, and no bonds; (c) high-valuation investors of type 11 carry equity, a negative bond position, and no money; (d) low-valuation investors of type 11, carry no equity, and are indifferent between carrying money or nonnegative bond holdings; (e) investors of type 01 carry the equity they had at the beginning of the period and are indifferent between borrowing, lending, or holding money, since the nominal interest rate on inside bonds in this case is $\rho^m = 0$.

In economy $\mathcal{E}^m_3$, investor's post trade portfolios in the OTC market are as follows: (a) high-valuation investors of type 10 carry equity, no money, and no bonds; (b) A fraction

$$\chi^L_{10} = \frac{(\alpha_{10}+\alpha_{11})(\alpha_{01}+\alpha_{11})G_L - \alpha_{11}(\alpha_{01}+\alpha_{10}+\alpha_{11})G_H\frac{\lambda}{1-\lambda}}{\alpha_{01}\alpha_{10}G_L} \tag{57}$$

of low-valuation investors of type 10 carry equity, no money, and no bonds, and the remaining $1-\chi^L_{10}$ fraction carry money, no equity, and no bonds; (c) high-valuation investors of type 11 carry equity, a negative bond position, and no money; (d) low-valuation investors of type 11, carry no equity, positive bond holdings, and no money; (e) investors of type 01 carry a positive bond position, no money, and the equity they had at the beginning of the period. The nominal interest rate on inside bonds in this case, $\rho^m = \frac{\varepsilon^*_{11}-\varepsilon_L}{\varphi}\rho$, is positive, but low enough such that all high-valuation investors of type 11 prefer shorting the bond rather than lending. Figure 2 illustrates the existence regions in the space of parameters ι (vertical axis) and λ (horizontal axis).[21] Figure 3 shows the special case with $\alpha_{01} = 0$.

4 Cashless limits

In this section we study the properties of the monetary equilibrium in situations where agents can economize in the use of cash, and in settings where aggregate real cash balances are small relative to the aggregate real value of the stock of finacial assets being traded. In other words, we study economies that can be construed as approximations to *pure-credit* or *cashless economies*.[22] To this end, we focus on the baseline economy with general G and $\alpha_{01} = 0$ and study the limiting equilibrium as either: (i) $\lambda \to 1$, (ii) $\iota \to \bar{\iota}(\lambda)$, or (iii) $\alpha_{10} \to 0$. The first limit approximates an economy where investors of type 11 can buy assets with zero margin (infinite leverage). Thus in this case a measure α_{11} of investors can greatly economize on cash use along the intensive margin of trade. The second limit operates directly on the opportunity cost of holding of money and approximates an economy where this cost is so high that aggregate real balances become negligible. The third limit approximates an economy where—regardless of the equilibrium value of money—nobody *needs* money to trade, in the sense that it is budget feasible for every individual investor to long equity by shorting the bond, without ever using money as means of payment. In each of these limits we study aggregate real balances, transaction velocity, and the role of monetary policy on trading activity and asset prices.

In a stationary equilibrium, we can define *transaction velocity* of money, $\mathcal{V}$, as the ratio of the nominal value of all transactions to the money supply, i.e., $\mathcal{V} \equiv \frac{[\alpha_{10}G(\varepsilon^*_{10})+\alpha_{11}G(\varepsilon^*_{11})]p_t A^s}{A^m_t} = [\alpha_{10}G(\varepsilon^*_{10}) + \alpha_{11}G(\varepsilon^*_{11})]\,(\varepsilon^*_{10}+\phi^s)/Z$. Transaction velocity is useful in this context as a measure of the "efficiency of cash use", i.e., as a measure of the average number of transactions that can

[21] It is easy to prove that $\hat{\iota}\,(0) < \hat{\iota}(\hat{\lambda}) \leq \bar{\iota}$ (the latter with "=" if $\alpha_{01} = 0$) and

$$\frac{\partial \hat{\iota}\,(\lambda)}{\partial \lambda} = \frac{\pi_H\,(\varepsilon_H - \varepsilon_L)\,\alpha_{11}\theta\,\{\pi_H\,(1-\alpha_{10}-\alpha_{11})\,\varepsilon_H + [1-\pi_H\,(1-\alpha_{10}-\alpha_{11})]\,\varepsilon_L\}}{[(1-\lambda)\,\bar{\varepsilon} + \alpha_{11}\theta\pi_H\lambda\,(\varepsilon_H - \varepsilon_L)]^2} > 0.$$

Notice that $\hat{\iota}\,(\lambda) = \bar{\iota}$ for all λ if $\theta = 0$. Also, notice that $\hat{\lambda} \to \pi_L$ as $\alpha_{01} \to 0$.

[22] In this paper we follow the tradition in monetay theory to use "cash" as a synonym of unbacked "money." Hence by "cashless economy" we mean a "moneyless economy" akin to the "pure credit system" envisioned by Wicksell (1898, p. 62-70) that Woodford (2003) uses to motivate his moneyless appraoch to monetary economics. In lay terms, in contrast, "cash" often refers to currency, so with this alternative terminology, it would be possible for a payment system to be cashless but not moneyless.

be supported per dollar in circulation.[23] In the limiting economy as $\Delta \to 0$, we have

$$\mathcal{V} \equiv [\alpha_{10}G(\varepsilon^*_{10}) + \alpha_{11}G(\varepsilon^*_{11})] \frac{\varphi}{\mathcal{Z}} \tag{58}$$
$$= \begin{cases} \frac{\{\alpha_{10}[1-G(\varepsilon^*_{10})]+\alpha_{11}\}[\alpha_{10}G(\varepsilon^*_{10})+\alpha_{11}\lambda]}{\alpha_{10}G(\varepsilon^*_{10})} & \text{if } \hat{\iota}(\lambda) < \iota < \bar{\iota}(\lambda) \\ \frac{(\alpha_{10}+\alpha_{11})(\alpha_{10}+\alpha_{11}\frac{1}{1-\lambda})G(\varepsilon^*)[1-G(\varepsilon^*)]}{\alpha_{10}G(\varepsilon^*)+\alpha_{11}\frac{1}{1-\lambda}[G(\varepsilon^*)-\lambda]} & \text{if } 0 < \iota \leq \hat{\iota}(\lambda). \end{cases}$$

Proposition 6 *Consider the limiting economy (as $\Delta \to 0$) with $\alpha_{10}, \alpha_{11} \in (0,1)$. Let*

$$\varphi^n_{\lambda=1} \equiv \lim_{\lambda \to 1} \varphi^n = \bar{\varepsilon} + \alpha_{11}\theta(\varepsilon_H - \bar{\varepsilon}).$$

As $\lambda \to 1$, $\varepsilon^*_{11} = \varepsilon^n \to \varepsilon_H$,

$$\frac{\mathcal{Z}}{\varphi} \to \frac{\alpha_{10}G(\varepsilon^*_{10})}{[1 - G(\varepsilon^*_{10})]\alpha_{10} + \alpha_{11}} \tag{59}$$

$$\mathcal{V} \to \frac{\{\alpha_{10}[1 - G(\varepsilon^*_{10})] + \alpha_{11}\}[\alpha_{10}G(\varepsilon^*_{10}) + \alpha_{11}]}{\alpha_{10}G(\varepsilon^*_{10})} \tag{60}$$

$$\varphi \to \varphi^n_{\lambda=1} + [\alpha_{10} + \alpha_{11}(1-\theta)] \int_{\varepsilon_L}^{\varepsilon^*_{10}} (\varepsilon^*_{10} - \varepsilon)\, dG(\varepsilon) \tag{61}$$

where $\varepsilon^*_{10} \in [\varepsilon_L, \varepsilon_H]$ is the unique solution to

$$\frac{\alpha_{11}\theta(\varepsilon_H - \varepsilon^*_{10}) + [\alpha_{10} + \alpha_{11}(1-\theta)] \int_{\varepsilon^*_{10}}^{\varepsilon_H} (\varepsilon - \varepsilon^*_{10})\, dG(\varepsilon)}{\bar{\varepsilon} + [\alpha_{10} + \alpha_{11}(1-\theta)] \int_{\varepsilon_L}^{\varepsilon^*_{10}} (\varepsilon^*_{10} - \varepsilon)\, dG(\varepsilon) + \alpha_{11}\theta(\varepsilon_H - \bar{\varepsilon})} = \iota. \tag{62}$$

Proposition 6 shows that, as $\lambda \to 1$, real balances remain positive and velocity remains bounded as long as $\alpha_{10} \in \mathbb{R}_{++}$ and $\iota < \bar{\iota}(1)$. This is not surprising given money demand in the OTC round is supported by low-valuation investors of type 10, and money demand in the second subperiod is supported by the probability of being a high-valuation investor of type 10 in the following period. The portfolio problem of the investor of type 11 remains well defined even as λ approaches 1 because the equilibrium real interest rate in that limit becomes very high, $\rho^s \to \frac{\varepsilon_H}{\varphi}\rho$, which in equilibrium limits their desire to short the bond.

[23] Money in our model may be traded many times in an OTC round. Consider a dollar initially spent by an equity buyer. If the dollar ends up in the hands of an equity seller of type 10, the dollar is not further circulated. But if the dollar ends up in the hands of an equity seller of type 11, it flows from the seller through the margin loans market to an equity buyer of type 11 who borrows. The dollar, now in the hands of the type 11 borrower, will travel further, to a seller and possibly further to another margin loan borrower, until it eventually ends up, unspent, in the post-OTC portfolio of a seller who carries it into the following subperiod (this seller is necessarily of type 10, if $q_t\phi^m_t < 1$ or possibly of type 11 if $q_t\phi^m_t = 1$). In this sense, we can think of margin loans as a liquidity saving mechanism, which helps sustain monetary equilibrium.

Proposition 7 *Consider the limiting economy (as $\Delta \to 0$) with $\alpha_{10}, \alpha_{11} \in (0,1)$. As $\iota \to \bar{\iota}(\lambda)$, $\varepsilon_{11}^* = \varepsilon^n$, $\varepsilon_{10}^* \to \varepsilon_L$,*

$$\frac{\mathcal{Z}}{\varphi} \to 0 \tag{63}$$

$$\mathcal{V} \to \infty \tag{64}$$

$$\varphi \to \varphi^n. \tag{65}$$

Proposition 7 is a familiar "cashless limit" in micro-founded monetary models: as the opportunity cost of holding money becomes very large, real balances approach zero. In this case, since the real value of the equity purchases of the α_{11} type 11 investors remains positive in the limit, the transaction velocity of money diverges to infinity (but the velocity for corresponding only to type 10 investors, remains bounded). To see this clearly, notice we could decompose $\mathcal{V}$ into two components, i.e., $\mathcal{V} = \mathcal{V}_{10} + \mathcal{V}_{11}$, where $\mathcal{V}_{10} \equiv \alpha_{10} G(\varepsilon_{10}^*) \frac{\varphi}{\mathcal{Z}}$ and $\mathcal{V}_{11} \equiv \alpha_{11} G(\varepsilon_{11}^*) \frac{\varphi}{\mathcal{Z}}$ are the contributions to velocity of transactions by agents of type 10, and agents of type 11, respectively. Explicitly,

$$\mathcal{V}_{10} = \begin{cases} \left[1 - G(\varepsilon_{10}^*)\right] \alpha_{10} + \alpha_{11} & \text{if } \hat{\iota}(\lambda) < \iota < \bar{\iota}(\lambda) \\ \frac{\alpha_{10} G(\varepsilon^*)[1-G(\varepsilon^*)]\left(\alpha_{10}+\alpha_{11}\frac{1}{1-\lambda}\right)}{\alpha_{10} G(\varepsilon^*) + \alpha_{11}\frac{1}{1-\lambda}[G(\varepsilon^*)-\lambda]} & \text{if } 0 < \iota \le \hat{\iota}(\lambda) \end{cases}$$

$$\mathcal{V}_{11} = \begin{cases} \frac{\alpha_{11}\lambda\left\{\left[1-G(\varepsilon_{10}^*)\right]\alpha_{10}+\alpha_{11}\right\}}{\alpha_{10} G(\varepsilon_{10}^*)} & \text{if } \hat{\iota}(\lambda) < \iota < \bar{\iota}(\lambda) \\ \frac{\alpha_{11} G(\varepsilon^*)[1-G(\varepsilon^*)]\left(\alpha_{10}+\alpha_{11}\frac{1}{1-\lambda}\right)}{\alpha_{10} G(\varepsilon^*) + \alpha_{11}\frac{1}{1-\lambda}[G(\varepsilon^*)-\lambda]} & \text{if } 0 < \iota \le \hat{\iota}(\lambda). \end{cases}$$

Hence as $\iota \to \bar{\iota}(\lambda)$, we have $\mathcal{V}_{11} \to \infty$ and $\mathcal{V}_{10} \to (\alpha_{10} + \alpha_{11})$.

For the following result, it is useful to define $\alpha_{10} \equiv \alpha_s \alpha$, $\alpha_{11} \equiv \alpha_s (1-\alpha)$, and $\alpha_{01} = 0$, with $\alpha_s \in (0,1]$, and $\alpha \in [0,1]$. Let $\bar{\varsigma}(\alpha)$ and $\hat{\varsigma}(\alpha)$ denote the same bounds defined in (41) and (42), but regarded as functions of α, i.e.,

$$\bar{\varsigma}(\alpha) \equiv \frac{(1-\alpha)\theta(\varepsilon^n - \varepsilon_L) + [\alpha + (1-\alpha)(1-\theta)](\bar{\varepsilon} - \varepsilon_L) + (1-\alpha)\theta\frac{1}{1-\lambda}\int_{\varepsilon^n}^{\varepsilon_H}(\varepsilon - \varepsilon^n)\,dG(\varepsilon)}{\frac{\bar{\varepsilon}}{\alpha_s} + (1-\alpha)\theta\left[\int_{\varepsilon_L}^{\varepsilon^n}(\varepsilon^n - \varepsilon)\,dG(\varepsilon) + \frac{\lambda}{1-\lambda}\int_{\varepsilon^n}^{\varepsilon_H}(\varepsilon - \varepsilon^n)\,dG(\varepsilon)\right]}$$

$$\hat{\varsigma}(\alpha) \equiv \frac{\left\{\alpha + (1-\alpha)\left[(1-\theta) + \theta\frac{1}{1-\lambda}\right]\right\}\int_{\varepsilon^n}^{\varepsilon_H}(\varepsilon - \varepsilon^n)\,dG(\varepsilon)}{\frac{\bar{\varepsilon}}{\alpha_s} + \int_{\varepsilon_L}^{\varepsilon^n}(\varepsilon^n - \varepsilon)\,dG(\varepsilon) + (1-\alpha)\theta\frac{\lambda}{1-\lambda}\int_{\varepsilon^n}^{\varepsilon_H}(\varepsilon - \varepsilon^n)\,dG(\varepsilon)}.$$

Proposition 8 *Consider the limiting economy (as $\Delta \to 0$) with $\alpha_{10} \equiv \alpha_s \alpha$, $\alpha_{11} \equiv \alpha_s(1-\alpha)$, $\alpha_s \in (0,1]$, $\alpha \in [0,1]$, and $\lambda \in (0,1]$. Let*

$$\tilde{\varphi}^n \equiv \lim_{\alpha \to 0} \varphi^n = \bar{\varepsilon} + \alpha_s \theta \left[\int_{\varepsilon_L}^{\varepsilon^n}(\varepsilon^n - \varepsilon)\,dG(\varepsilon) + \frac{\lambda}{1-\lambda}\int_{\varepsilon^n}^{\varepsilon_H}(\varepsilon - \varepsilon^n)\,dG(\varepsilon)\right]. \tag{66}$$

As $\alpha \to 0$*,*

(i) If $\hat{\varsigma}(0) < \iota < \bar{\varsigma}(0)$*, then*

$$\frac{\mathcal{Z}}{\varphi} \to 0 \tag{67}$$

$$\mathcal{V} \to \infty \tag{68}$$

$$\varphi \to \tilde{\varphi} \equiv \tilde{\varphi}^n + \alpha_s (1-\theta) \int_{\varepsilon_L}^{\varepsilon_{10}^*} (\varepsilon_{10}^* - \varepsilon)\, dG(\varepsilon), \tag{69}$$

where $\varepsilon_{10}^* \in (\varepsilon_L, \varepsilon^n)$ *is the unique solution to*

$$\frac{\theta\left(\varepsilon^n - \varepsilon_{10}^*\right) + (1-\theta)\int_{\varepsilon_{10}^*}^{\varepsilon_H}\left(\varepsilon - \varepsilon_{10}^*\right) dG(\varepsilon) + \theta\frac{1}{1-\lambda}\int_{\varepsilon^n}^{\varepsilon_H}(\varepsilon - \varepsilon^n)\, dG(\varepsilon)}{\frac{\bar{\varepsilon}}{\alpha_s} + (1-\theta)\int_{\varepsilon_L}^{\varepsilon_{10}^*}\left(\varepsilon_{10}^* - \varepsilon\right) dG(\varepsilon) + \theta\left[\int_{\varepsilon_L}^{\varepsilon^n}(\varepsilon^n - \varepsilon)\, dG(\varepsilon) + \frac{\lambda}{1-\lambda}\int_{\varepsilon^n}^{\varepsilon_H}(\varepsilon - \varepsilon^n)\, dG(\varepsilon)\right]} = \iota. \tag{70}$$

(ii) If $0 < \iota \leq \hat{\varsigma}(0)$*, then*

$$\frac{\mathcal{Z}}{\varphi} \to \frac{G(\varepsilon^*) - \lambda}{1 - G(\varepsilon^*)} \tag{71}$$

$$\mathcal{V} \to \frac{\alpha_s G(\varepsilon^*)\left[1 - G(\varepsilon^*)\right]}{G(\varepsilon^*) - \lambda} \tag{72}$$

$$\varphi \to \tilde{\varphi} \equiv \bar{\varepsilon} + \alpha_s \left[\int_{\varepsilon_L}^{\varepsilon^*} (\varepsilon^* - \varepsilon)\, dG(\varepsilon) + \theta \frac{\lambda}{1-\lambda} \int_{\varepsilon^*}^{\varepsilon_H} (\varepsilon - \varepsilon^*)\, dG(\varepsilon)\right], \tag{73}$$

where $\varepsilon^* \in [\varepsilon^n, \varepsilon_H)$ *is the unique solution to*

$$\frac{\left(1 - \theta + \theta\frac{1}{1-\lambda}\right)\int_{\varepsilon^*}^{\varepsilon_H}(\varepsilon - \varepsilon^*)\, dG(\varepsilon)}{\frac{\bar{\varepsilon}}{\alpha_s} + \int_{\varepsilon_L}^{\varepsilon^*}(\varepsilon^* - \varepsilon)\, dG(\varepsilon) + \theta\frac{\lambda}{1-\lambda}\int_{\varepsilon^*}^{\varepsilon_H}(\varepsilon - \varepsilon^*)\, dG(\varepsilon)} = \iota. \tag{74}$$

Proposition 8 considers the limiting economy as the fraction of investors who do not have access to margin loans vanishes, while keeping constant the proportion of investors who can trade equity. In other words, as $\alpha \to 0$, $\alpha_{10} \to 0$ and $\alpha_{11} \to \alpha_s$. This limiting economy is one where it is budget feasible for virtually every investor to finance all their equity purchases by shorting the bond, i.e., investors can purchase equity shares even if they come into the period carrying no cash. As explained in the context of Proposition 2, in a monetary equilibrium, investors of type 10 with relatively low valuation for equity are the ones who always demand money in the OTC trading round, while investors of type 11 with relatively low valuation demand money in the OTC round only if $0 < \iota \leq \hat{\varsigma}(0)$. Therefore the first-order implication of letting α_{10} approach zero is that the extensive margin of money demand from investors of type 10, i.e., the number of type 10 investors who wish to hold money, approaches zero.[24]

If the nominal policy rate is relatively low, i.e., if $0 < \iota \leq \hat{\varsigma}(0)$ as in part (*ii*) of Proposition 8, then the aggregate money demand from type 10 investors vanishes in the limit, but the aggregate money demand from type 11 investors with low valuation remains positive in the limit. The reason why money can have value in this limiting economy can be understood at least from two perspectives. First, money demand can be positive overnight because the collateral constraint is expected to be binding for type 11 investors, so money can allow them to take a long position in equity that is larger than the one they would be able to take relying only on the margin loan.[25] Second, low-valuation investors of type 11 are willing to hold cash at the end of the OTC round because the nominal rate on the bond (i.e., the opportunity cost of holding cash in the OTC round) is zero when $0 < \iota \leq \hat{\varsigma}(0)$; this makes them indifferent between holding their wealth in money or bonds, and they hold some of each. As a result, in economies that satisfy $0 < \iota \leq \hat{\varsigma}(0)$, i.e., economies with relatively low inflation and relatively low ability to take on leverage, real balances converge to a postive limit (i.e., (71)), and velocity converges (i.e., (72)) as $\alpha_{10} \to 0$.

If the nominal policy rate is relatively high, i.e., if $\hat{\varsigma}(0) < \iota < \bar{\varsigma}(0)$ as in part (*i*) of Proposition 8, then real balances converge to zero (i.e., (67)), and transaction velocity goes to infinity (i.e., (68)) as $\alpha_{10} \to 0$. The immediate reason for this result is that given this high policy rate, total money demand in the OTC round vanishes as $\alpha_{01} \to 0$, for two reasons. First, low-valuation investors of type 10 would be willing to hold money, but the number of type 10 investors vanishes as $\alpha_{10} \to 0$. Second, low-valuation investors of type 11 are unwilling to hold money because money is dominated in rate of return by the collateralized inside bond. Since virtually nobody wishes to hold money in the OTC round, money has no value in the limiting

[24]The restriction that $\lambda \neq 0$ in the statement of Proposition 8 is because, as shown in Figure 1, $\lambda = 0$ implies $\hat{\varsigma}(\alpha) = \bar{\varsigma}(\alpha)$ for all α, and therefore in that case monetary equilibrium exists only for $\iota = \hat{\varsigma}(\alpha) = \bar{\varsigma}(\alpha)$. From Figure 1 we can also see that $\lambda = 1$ implies $\hat{\varsigma}(\alpha) = 0$ for all α, so low-valuation investors of type 11 hold money in the OTC round only if $\iota = \hat{\varsigma}(\alpha) = 0$. Thus if $\lambda = 1$, the range in part (*ii*) of Proposition 8 collapses to a single point, and only the parameter region in part (*i*) is generic in that case.

[25]Notice that in the range $0 < \iota \leq \hat{\varsigma}(0)$, this argument requires $\lambda < 1$. The reason is that from Figure 1 we can see that $\lambda = 1$ implies $\hat{\varsigma}(\alpha) = 0$ for all α, so low-valuation investors of type 11 hold money in the OTC round only if $\iota = \hat{\varsigma}(\alpha) = 0$. In other words, if $\lambda = 1$ then the range in part (*ii*) of Proposition 8 collapses to a single point (money is only held exactly at the Friedman rule), and only the parameter region in part (*i*) is generic in that case.

economy $\alpha_{10} \to 0$ with $\iota \in (\hat{\varsigma}(0), \bar{\varsigma}(0))$. The key result is that

$$\lim_{\alpha \to 0} \frac{\mathcal{Z}}{\varphi} = \lim_{\alpha \to 0} \frac{1}{\mathcal{V}} = 0 < \lim_{\alpha \to 0} (\varphi - \varphi^n) = \alpha_s (1 - \theta) \int_{\varepsilon_L}^{\varepsilon_{10}^*} (\varepsilon_{10}^* - \varepsilon) \, dG(\varepsilon). \tag{75}$$

Notably, even though real balances and velocity converge to their nonmonetary equilibrium levels as $\alpha_{10} \to 0$, the real equity price in this cashless limit of the monetary economy exceeds the (corresponding limit of the) nonmonetary-equilibrium price by the value of a resale-option term in (75). Since ε_{10}^* is a function of ι (from (70)), it follows that in the cashless limit the asset price is still responsive to monetary policy, and that the magnitude of this response remains bounded away from zero even though the quantity of real balances converges to zero. This result stands in contrast with the more conventional result that the monetary equilibrium prices and allocations converge to their nonmonetary equilibrium counterparts when real balances vanish, e.g., as it happens with the cashless limit in Proposition 7.

Why does the cashless limit in Proposition 8 not correspond to the nonmonnetary equilibrium? The answer is in (75) and, at least algebraically, it is simple: because the following two conditions are met: (a) the limit, as $\alpha \to 0$, of the marginal valuation of investors of type 10 is strictly larger than ε_L (i.e., the ε_{10}^* that solves (70) satisfies $\varepsilon_L < \varepsilon_{10}^*$, given that $\hat{\varsigma}(0) < \iota < \bar{\varsigma}(0)$), and (b) $\theta < 1$. Condition (a) is a result of the fact that the asset demand from leveraged high-valuation investors of type 11 sustains an equilibrium asset price in the limit that is high enough to encourage investors of type 10 with low enough valuations, to sell the asset for cash at that price. Condition (b) is a parametric condition on the market power of financial intermediaries. The case with $\theta = 1$ correponds to a market structure where bond brokers have no market power to extract intermediation fees, and in this case the equity price in the cashless limit corresponds to the asset price in the nonmonetary equilibrium. To understand this result at a coneptual level, it is useful to review the components of the limiting equity price

$$\lim_{\alpha \to 0} \varphi = \bar{\varepsilon} + \alpha_s \theta \frac{\lambda}{1 - \lambda} \int_{\varepsilon^n}^{\varepsilon_H} (\varepsilon - \varepsilon^n) \, dG(\varepsilon)$$
$$+ \alpha_s \theta \int_{\varepsilon_L}^{\varepsilon^n} (\varepsilon^n - \varepsilon) \, dG(\varepsilon) + \alpha_s (1 - \theta) \int_{\varepsilon_L}^{\varepsilon_{10}^*} (\varepsilon_{10}^* - \varepsilon) \, dG(\varepsilon).$$

The first term is the expected value of the dividend flow. The second term is the investor's share θ of the expected value of the pledge option on the asset (i.e., the marginal value to the investor of the asset when it is pledged as collateral for shorting the bond). The third term is the investor's share θ of the expected value of the resale option on the asset. The sum of the first three terms equals φ_0^n (the nonmonetary equity price in the economy with $\alpha = 0$). The fourth term is the key term since it is the component of the asset price that is present in the cashless limit but not in the nonmonetary equilibrium. This term captures the improvement in the investor's bargaining position with the bond dealer due to the fact that the investor's outside option is not just autarky, but rather trading in the equity market without access to bonds. To see this clearly, consider an investor of type 11 who draws $\varepsilon \in [\varepsilon_L, \varepsilon^n]$. He wishes to *sell* all his equity and simultaneously bargains with a bond broker to take a long position in the bond. The investor's outside option in this bargain is the payoff of trading in the equity market without access to the bond market. The particular kind of trade that this outside trading option would entail, depends on the investor's valuation. If $\varepsilon \in [\varepsilon_L, \varepsilon_{10}^*]$, then the investor's outside option amounts to acting like an investor of type 10, i.e., *selling* equity and holding the cash. The expected capital gain from reselling the asset is $\int_{\varepsilon_L}^{\varepsilon_{10}^*} (\varepsilon_{10}^* - \varepsilon)\, dG(\varepsilon)$, and the bargaining outcome assigns a fraction $1 - \theta$ of this gain from trade to the investor. This is precisely the right side of (75), i.e., the additional term in the equity price at the cashless limit.[26]

Intuitively, the reason for (75) is that an investor of type 11 with valuation $\varepsilon \in [\varepsilon_L, \varepsilon_{10}^*]$ has the option to sell equity for cash in the equity market, and this option improves his bargaining position with bond dealers, and allows the investor to earn a larger share of gain from his portfolio reallocation (from equity to bonds). The key result is then that the value of this option to the individual investor remains strictly positive in the limit as $\alpha \to 0$, even as the aggregate real money balances are converging to zero along with the volume of cash transactions in the equity market. This may seem counterintuitive: why isn't the value of the investor's outside option vanishing in the limit? There are several ways to bring the equilibirum conditions to clarify these kinds of questions. First, it may help to notice that the trade each individual investor of type 11 may want to execute in the equity market is infinitesimal, so the investor of type 11 can always execute it no matter how small the equity market has gotten along the trajectory toward the cashless limit. The second question that may arise is: if the type 11 investor with valuation $\varepsilon \in [\varepsilon_L, \varepsilon_{10}^*]$ where to execute his option to sell his portfolio for cash in the equity market along the trajectory toward the cashless limit (he actually does not exercise

[26]If $\varepsilon \in [\varepsilon_{10}^*, \varepsilon^n]$, then the investor's outside option again amounts to acting like an investor of type 10, which in this case entails *buying* equity with cash. Hence, since the equity would not be sold for this range of valuations, it would not deliver a gain associated with reselling, and therefore the investor's outside option of trading in the equity market with no access to bonds does not appear in the expected resale value component of the asset. Algebraically, this is why the upper bound of the integral (75) is ε_{10}^* rather than ε^n.

this option in the equilibrium), who would be on the other side of this transaction buying equity for cash? The answer is: investors of type 11 with valuations in the interval $[\varepsilon_{11}^*, \varepsilon_H]$, along with the *very few* investors of type 10 with valuations in the interval $[\varepsilon_{10}^*, \varepsilon_H]$ that are still around along the trajectory toward the cashless limit. Notice that the latter may be few, but each is holding strictly positive real balances along the cashless trajectory as $\alpha \to 0$. That is, although the aggregate real balances are converging to zero (as shown by (67)), the limit of the *real balances per investor of type* 10 is strictly positive, i.e.,

$$\lim_{\alpha \to 0} \frac{\mathcal{Z}/\varphi}{\alpha_{10}} = \lim_{\alpha \to 0} \frac{G(\varepsilon_{10}^*)}{[1 - G(\varepsilon_{10}^*)]\alpha_{10} + \alpha_{11}} = \frac{G(\varepsilon_{10}^*)}{\alpha_{11}} > 0.$$

This is a key difference with the conventional inflationary cashless limit of Proposition 7. In that case, $\lim_{\iota \to \bar{\iota}(\lambda)} \varepsilon_{10}^* = \varepsilon_L$, and therefore

$$\lim_{\iota \to \bar{\iota}(\lambda)} \frac{\mathcal{Z}/\varphi}{\alpha_{10}} = \lim_{\iota \to \bar{\iota}(\lambda)} \frac{G(\varepsilon_{10}^*)}{[1 - G(\varepsilon_{10}^*)]\alpha_{10} + \alpha_{11}} = 0,$$

which is why the value of the option to sell equity for cash in the equity market (for a low-valuation investor of type 11) converges to zero in the cashless limit of Proposition 7, and consequently the equity price in the cashless limit is equal to the equity price in the nonmonetary equilibrium.

To summarize, the equilibrium in the cashless limit as $\alpha \to 0$ behaves like this: the resale value option in (75) is positive because even in the limit, and individual investor of type 11 can threaten to sell his equity for cash in the equity market, and this off-equilibrium threat improves his bargaining position with the bond broker allowing the investor to earn a larger share of the trade surplus. Since the resale value option is positive in the limit, the asset price remains relatively high, and in particular, higher than the price in the nonmonetary equilibrium (where, with no cash valued in equilibrium, the investor cannot threaten to sell equity for cash in the equity market to improve his bargaining position). In other words, there is a discontinuity in the equilibrium value of this outside option. On the one hand, this outside option is absent in the nonmonetary equilibrium. On the other hand, in the limit as $\alpha \to 0$, aggregate real balances converge to 0 but the limit of the value of the outside option to sell equity for cash is positive for any investor of type 11. Thus the resale value option of the type 11 in (75) is *discontinuous* when taking the limit with respect to α_{10}, and the asset price inherits this discontinuity. In turn, the relatively high equity price in the cashless limit rationalizes the fact that not all type 10 investors prefer buying equity rather than holding cash in the OTC, which manifiests itself as $\varepsilon_{10}^* > \varepsilon_L$ in the cashless limit, and therefore $\lim_{\alpha \to 0} \frac{\mathcal{Z}/\varphi}{\alpha_{10}} > 0$ in the cashless limit.

To conclude, notice that the result (75) depends on two fundamental features of the environmnent. First, as explained above, bond dealers must have at least some degree of market power, i.e., we must have $\theta < 1$. Second, investors must have at least some ability to take on leverage, i.e., we must have $\lambda > 0$. This is so because $\hat{\varsigma}(0) \leq \bar{\varsigma}(0)$ for all $\lambda \in [0,1]$, but with "=" only if $\lambda = 0$, so the parameter region in part (i) of Proposition 8 collapses to a single point if $\lambda = 0$. But for any positive λ, no matter how small, the region of discontinuity has positive measure in parameter space.

4.1 Example

In this section we consider the cashless limits for the special case $G(\varepsilon) = G_L \mathbb{I}_{\{\varepsilon_L \leq \varepsilon < \varepsilon_H\}} + \mathbb{I}_{\{\varepsilon_H \leq \varepsilon\}}$, for which equilibrium can be solved for in closed form. For the following result, it is useful to define $\alpha_{10} \equiv \alpha_s \alpha$, $\alpha_{11} \equiv \alpha_s (1-\alpha)$, and $\alpha_{01} = 0$, with $\alpha_s \in (0,1]$, and $\alpha \in [0,1]$. Let $\bar{\varsigma}(\alpha)$ and $\hat{\varsigma}(\alpha)$ denote the same bounds defined in (45) and (46), but regarded as functions of α, i.e.,

$$\bar{\varsigma}(\alpha) \equiv \frac{[(1-\alpha) G_L \theta + G_H](\varepsilon_H - \varepsilon_L)}{\frac{\bar{\varepsilon}}{\alpha_s} + (1-\alpha)\theta G_L (\varepsilon_H - \varepsilon_L)}$$

$$\hat{\varsigma}(\alpha) \equiv \frac{\left[\alpha + (1-\alpha)\left(1 - \theta + \theta \frac{1}{1-\lambda}\right)\right] G_H (\varepsilon_H - \varepsilon_L)}{\frac{\bar{\varepsilon}}{\alpha_s} + (1-\alpha)\theta G_H \frac{\lambda}{1-\lambda}(\varepsilon_H - \varepsilon_L)}.$$

Also, assume $\alpha_{01} = 0$, so (47) simplifies to $\hat{\lambda} = G_L$. In this case,

$$\lim_{\alpha \to 0} \varphi^n = \begin{cases} \bar{\varepsilon} + \alpha_s \theta G_H (\varepsilon_H - \varepsilon_L) \frac{\lambda}{1-\lambda} & \lambda < G_L \\ \bar{\varepsilon} + \alpha_s \theta G_L (\varepsilon_H - \varepsilon_L) & G_L \leq \lambda. \end{cases}$$

Proposition 9 *Consider the limiting economy (as $\Delta \to 0$) with $\alpha_{10} \equiv \alpha_s \alpha$, $\alpha_{11} \equiv \alpha_s (1-\alpha)$, $\alpha_s \in (0,1]$, $\alpha \in [0,1]$, and $\lambda \in [0,1]$. As $\alpha \to 0$,*

(i) If $\iota < \bar{\varsigma}(0)$ and $G_L < \lambda$ then

$$\varepsilon_L < \lim_{\alpha \to 0} \varepsilon_{10}^* = \varepsilon_L + \frac{\bar{\varsigma}(0) - \iota}{\frac{\bar{\varsigma}(0)}{\varepsilon_H - \varepsilon_L} + \iota \frac{[\alpha_{10} + \alpha_{11}(1-\theta)] G_L}{\bar{\varepsilon} + \alpha_{11}\theta G_L (\varepsilon_H - \varepsilon_L)}} < \varepsilon_{11}^* = \varepsilon_H$$

and

$$\frac{\mathcal{Z}}{\varphi} \to 0 \tag{76}$$

$$\mathcal{V} \to \infty \tag{77}$$

$$\varphi \to \frac{\bar{\varepsilon} + \alpha_s G_L (\varepsilon_H - \varepsilon_L)}{1 + \iota \frac{(1-\theta) G_L}{\theta + (1-\theta) G_H}}. \tag{78}$$

(ii) If $0 < \iota < \hat{\varsigma}(0)$ *and* $\lambda < G_L$*, then*

$$\varepsilon_L < \lim_{\alpha \to 0} \varepsilon^* = \varepsilon_H - \frac{\iota[\bar{\varepsilon} + \alpha_s G_L(\varepsilon_H - \varepsilon_L)]}{\left[\alpha_s\left(1+\theta\frac{\lambda}{1-\lambda}\right)\right]G_H + \iota\alpha_s\left[1 - G_H\left(1+\theta\frac{\lambda}{1-\lambda}\right)\right]} < \varepsilon_H$$

and

$$\frac{\mathcal{Z}}{\varphi} \to \frac{G_L - \lambda}{G_H} \tag{79}$$

$$\mathcal{V} \to \frac{\alpha_s G_L G_H}{G_L - \lambda} \tag{80}$$

$$\varphi \to \frac{\bar{\varepsilon} + \alpha_s G_L(\varepsilon_H - \varepsilon_L)}{1 + \iota\frac{\left[1 - G_H\left(1+\theta\frac{\lambda}{1-\lambda}\right)\right]}{G_H\left(1+\theta\frac{\lambda}{1-\lambda}\right)}}. \tag{81}$$

With (78) it is easy to show that $\lim_{\alpha \to 0}(\varphi - \varphi^n) > 0$ for all $\iota < \bar{\varsigma}(0)$. Similarly, (81) implies $\lim_{\alpha \to 0}(\varphi - \varphi^n) \geq 0$ for all $\iota \leq \hat{\varsigma}(0)$ (with "=" only if $\iota = \hat{\varsigma}(0)$). This example makes clear that $\alpha \to 0$ implies that the monetary equilibrium disappears in the cashless limit (in the sense that the value of money converges to zero) only if investors can take on sufficient leverage. If the borrowing constraints are relatively severe (e.g., $\lambda < G_L$ as in part (*ii*) of Proposition 9), then the monetary equilibrium survives as $\alpha \to 0$, in the sense that real balances approach a strictly positive limit.

5 Extensions

In this section we consider two extensions of the baseline model. The first generalizes the model to allow for capital accumulation. This extension allows us to explore the transmission of monetary policy to asset prices, and in turn, the transmission from asset prices to the real economy. The second extension verifies the robustness of the cashless limiting results under an alternative credit arrangement where, rather than having to use the asset as collateral, investors are able to issue unsecured debt up to a given limit. As with the baseline model, we formulate these extensions in discrete time, and consider the continuous-time approximation to characterize equilibrium.

5.1 Capital accumulation

In the baseline model, the number of production units, A^s, is exogenous and constant. This means that monetary policy and the details of the OTC marketstructure affect asset prices but do not affect conventional measures of real economic activity, such as aggregate output or investment. In this section we endogenize the productive capacity of the economy by letting agents invest to augment the stock of productive units.

The model is as in Section 2, with the following change. We regard the productive units that yield the dividend good as a *capital stock* that can be accumulated. Specifically, in the second subperiod of period t, investors have access to a production technology that transforms $n \in \mathbb{R}_+$ units of the general good into x units of capital according to $x = f_t(n)$, where the production function f_t is strictly increasing, differentiable, and satisfies $f_t(0) = \lim_{n\to\infty} f_t'(n) = 0$, and $f_t'(0) = \infty$. Thus, the value of an investor in the second subperiod is

$$W_t^I(\boldsymbol{a}_t, a_t^b, k_t) = \max_{(c_t, h_{1t}, h_{2t}, x_t, \tilde{\boldsymbol{a}}_{t+1}) \in \mathbb{R}_+^6} \left[c_t - h_t + \beta \mathbb{E}_t \int V_{t+1}^I(\boldsymbol{a}_{t+1}, \varepsilon)\, dG(\varepsilon) \right]$$

$$\text{s.t. } c_t + \boldsymbol{\phi}_t \tilde{\boldsymbol{a}}_{t+1} \le h_{1t} + \boldsymbol{\phi}_t \boldsymbol{a}_t + a_t^b - k_t + \phi_t^s x_t + T_t,$$

with $\boldsymbol{a}_{t+1} = (\tilde{a}_{t+1}^m, \eta \tilde{a}_{t+1}^s)$, $h_t = h_{1t} + h_{2t}$, and $x_t = f_t(h_{2t})$, where h_t is the labor input (effort) devoted to production of general goods (equal to the quantity of general goods produced), h_{1t} is the quantity of general goods used for consumption, and h_{2t} is the quantity of general goods used as input to produce new capital, x_t. This problem can be written as

$$\begin{aligned} W_t^I(\boldsymbol{a}_t, a_t^b, k_t) = {} & \boldsymbol{\phi}_t \boldsymbol{a}_t + a_t^b - k_t + T_t \\ & + \max_{\tilde{\boldsymbol{a}}_{t+1} \in \mathbb{R}_+^2} \left[-\boldsymbol{\phi}_t \tilde{\boldsymbol{a}}_{t+1} + \beta \mathbb{E}_t \int V_{t+1}^I(\boldsymbol{a}_{t+1}, \varepsilon)\, dG(\varepsilon) \right] \\ & + \max_{h_{2t} \in \mathbb{R}_+} \left[\phi_t^s f_t(h_{2t}) - h_{2t} \right]. \end{aligned}$$

This value function is the same as (8), except for the addition of the last term that represents the investor's profit from producing and selling new capital at the market price ϕ_t^s. The optimal quantity of general goods that the investor devotes to the production of capital goods, i.e., the h_{2t} that satisfies $\phi_t^s f_t'(h_{2t}) = 1$, is denoted $g_t(\phi_t^s)$, i.e.,

$$g_t(\phi_t^s) \equiv f_t'^{-1}(1/\phi_t^s). \tag{82}$$

The quantity of new capital created by an individual investor is $x_t(\phi_t^s) \equiv f_t(g_t(\phi_t^s))$. We can regard $x_t(\phi_t^s)$ as an individual investor's contribution to aggregate investment; aggregate capital investment is

$$X_t(\phi_t^s) \equiv x_t(\phi_t^s) N_I. \tag{83}$$

The assumptions on f imply aggregate investment is increasing in the market price of the equity shares of capital, i.e.,

$$X_t'(\phi_t^s) = -\frac{f_t'(g_t(\phi_t^s))}{(\phi_t^s)^2 f_t''(g_t(\phi_t^s))} N_I > 0.$$

The law of motion of the aggregate capital stock is

$$A_{t+1}^s = \eta (A_t^s + X_t), \tag{84}$$

where X_t is aggregate investment added to the capital stock at the end of period t, and $A_0^s \in \mathbb{R}_+$ is given.[27]

The definition of equilibrium for the economy with capital accumulation is the same as Definition 1, but with two additional equilibrium variables, namely $\{X_t, A_{t+1}^s\}_{t=0}^{\infty}$, and two additional equilibrium conditions, namely $X_t = X_t(\phi_t^s)$ and (84). A RNE is a nonmonetary equilibrium with the structure described in Definition 2. A RME is a monetary equilibrium in which: (i) real equity prices (general goods per equity share) are time-invariant linear functions of the aggregate dividend, i.e., $\phi_t^s = \phi^s y_t$, $p_t \phi_t^m \equiv \bar{\phi}_{10t}^s = \bar{\phi}_{10}^s y_t$, and $p_t/q_t \equiv \bar{\phi}_{11t}^s = \bar{\phi}_{11}^s y_t$ for some $\phi^s, \bar{\phi}_{10}^s, \bar{\phi}_{11}^s \in \mathbb{R}_+$; and ($ii$) real money balances are a constant proportion of aggregate output, i.e., $\phi_t^m A_t^m = Z A_t^s y_t$ for some $Z \in \mathbb{R}_{++}$. Hence in a RME, $\varepsilon_{10t}^* = (p_t \phi_t^m - \phi_t^s) \frac{1}{y_t} = \bar{\phi}_{10}^s - \phi^s \equiv \varepsilon_{10}^*$, $\varepsilon_{11t}^* = (p_t/q_t - \phi_t^s) \frac{1}{y_t} = \bar{\phi}_{11}^s - \phi^s \equiv \varepsilon_{11}^*$, and nominal prices are

$$p_t = (\varepsilon_{10}^* + \phi^s) \frac{A_t^m}{Z A_t^s} \tag{85}$$

$$\phi_t^m = \frac{Z A_t^s y_t}{A_t^m} \tag{86}$$

$$q_t = \frac{\varepsilon_{10}^* + \phi^s}{\varepsilon_{11}^* + \phi^s} \frac{A_t^m}{Z A_t^s y_t}. \tag{87}$$

For the analysis that follows, we generalize the money supply process of Section 2 to the following money-growth rule

$$\frac{A_{t+1}^m}{A_t^m} = \frac{A_{t+1}^s}{A_t^s} \mu. \tag{88}$$

Notice that just as in the model of Section 2, this monetary policy rule implies the gross inflation rate (as measured by the growth in the nominal price of equity shares) is constant and equal to μ, i.e., $p_{t+1}/p_t = \mu$. In the special case with $A_t^s = A^s$, (88) reduces to $A_{t+1}^m / A_t^m = \mu$, namely the money growth process in our baseline model.

[27]Since agents can now augment the stock of productive units, the beginning-of-period exogenous lump-sum endowment is no longer needed to offset the depreciation in the aggregate capital stock due to the idiosyncratic obsolesence shock that affects each individual unit of capital.

In a recursive equilibrium (monetary or nonmonetary), once the asset price ϕ^s has been found, aggregate investment is given by $X_t = X_t(\phi^s y_t)$, and the aggregate capital stock follows the stochastic difference equation $A^s_{t+1} = \eta\left[A^s_t + X_t(\phi^s y_t)\right]$. If the equilibrium is monetary, once $(\varepsilon^*_{10}, \varepsilon^*_{11}, \phi^s, Z)$ have been found, the implied equilibrium stochastic processes for the nominal prices, $\{p_t, \phi^m_t, q_t\}$, are given by (85), (86), and (87). Thus along a RME, $(\varepsilon^*_{10}, \varepsilon^*_{11}, \phi^s, Z, \chi_{11})$ are constant, while nominal prices $\{p_t, \phi^m_t, q_t\}$ are random variables whose evolutions over time are driven by the stochastic dividend process $\{y_t\}$, and possibly also by transitional dynamics.[28] To streamline the presentation, we assume

$$f_t(n) = \varpi_t n^\sigma, \text{ with } \sigma \in (0,1), \text{ and } \varpi_t \equiv (\sigma y_t)^{-\sigma}. \tag{89}$$

Let $X_t(\Delta)$ denote aggregate investment in the recursive equilibrium of the discrete-time economy where the length of the time period is Δ, and let $\mathcal{X}_t \equiv \lim_{\Delta\to 0} \frac{1}{\Delta} X_t(\Delta)$ (i.e., $\mathcal{X}_t$ is the *investment rate*). Then as $\Delta \to 0$, (84) can be approximated by

$$\dot{A}^s_t = \mathcal{X}_t - \delta A^s_t. \tag{90}$$

Next, we characterize the RNE and RME for the limiting economy as $\Delta \to 0$.

Proposition 10 *Consider the limiting economy (as $\Delta \to 0$) with capital accumulation, and $\alpha_{10}, \alpha_{11} \in (0,1)\,\mathbb{R}_{++}$.*

(i) There exists a unique recursive nonmonetary equilibrium, $(\varepsilon^n, \varphi^n, \mathcal{X}^n)$. Moreover, $(\varepsilon^n, \varphi^n)$ are as described in Proposition 1, aggregate investment rate is $\mathcal{X}^n = (\varphi^n/\rho)^{\frac{\sigma}{1-\sigma}} N_I$, and the capital stock follows (90) with $\mathcal{X}_t = \mathcal{X}^n$.

(ii) If $0 \le \iota < \bar{\iota}(\lambda)$, there exists a unique recursive monetary equilibrium, $(\varepsilon^*_{10}, \varepsilon^*_{11}, \chi, \varphi, \mathcal{Z}, \mathcal{X})$. Moreover, $(\varepsilon^*_{10}, \varepsilon^*_{11}, \chi, \varphi, \mathcal{Z})$ are as described in Proposition 2, aggregate investment rate is $\mathcal{X} = (\varphi/\rho)^{\frac{\sigma}{1-\sigma}} N_I$, and the capital stock follows (90) with $\mathcal{X}_t = \mathcal{X}$.

Proposition 10 delivers a link between the asset price, i.e., the relative price of capital in terms of consumption goods (φ^n in the RNE or φ in the RME), and aggregate investment.

[28] By way of example, notice that if $f_t = f$ and $y_t = y$ for all t, then (82) implies aggregate investment is constant, i.e., $X_t(\phi^s y_t) = X(\phi^s y)$ for all t, and A^s_{t+1} converges monotonically to the unique steady state $\bar{A}^s = \frac{\eta}{1-\eta} X(\phi^s y)$ from any initial condition A^s_0. Given the deterministic transition path $A^s_{t+1} = \eta\left[A^s_t + X(\phi^s y)\right]$, the money supply process $\{A^m_t\}$, and nominal prices $\{p_t, \phi^m_t, q_t\}$, just follow (88), (85), (86), and (87).

5.2 Unsecured credit

In our baseline formulation, we modeled credit in the form of margin loans mainly because it is the most common form of credit used in financial markets. In this section we verify the robustness of our main results to an alternative credit arrangement where in the OTC round, investors are able to issue unsecured debt up to a given limit. The only relevant difference in the model is that the last constraint in the bargaining problem (5) is replaced by

$$-\bar{B}_t \leq \overline{a}_t^b, \tag{91}$$

where $\bar{B}_t \geq 0$ is the credit limit faced by an individual agent in the OTC round of period t. Suppose a broker extends an investor a loan of L dollars in order to purchase A dollars worth of an asset. Suppose, as will be the case in the model, that the investor chooses to borrow the maximum amount possible, i.e., $L = \bar{B}_t$. In this case, the margin is $\mathcal{M} = 1 - \bar{B}_t/A$, leverage is $\mathcal{L} = A/\left(A - \bar{B}_t\right)$, and the loan-to-value ratio is $\mathcal{R} = \bar{B}_t/A$.

We focus on recursive equilibria. To this end, let

$$\bar{B}_t \equiv \Lambda \frac{(p_t/q_t)\, A^s}{N_I} \tag{92}$$

for some $\Lambda > 0$. In a nonmonetary economy, $\bar{\phi}_t^s \equiv p_t/q_t$, and therefore (92) amounts to assuming $\bar{B}_t \equiv \Lambda \frac{\bar{\phi}_t^s A^s}{N_I}$. Formulation (92) corresponds to an economy (monetary or nonmonetary) where the aggregate real borrowing capacity of investors expressed in terms of general goods, i.e., $N_I \bar{B}_t$, is a multiple Λ of the real value (expressed in terms of general goods) of the equity shares outstanding, $(p_t/q_t)\, A^s$.

The structure of the recursive equilibrium is as described in Definition 2 and Definition 3. We again consider the limiting economy as $\Delta \to 0$, and as before, let $\varphi \equiv \rho\phi^s$, $\mathcal{Z} \equiv \rho Z$, and $\iota \equiv i^p/\rho$. For the following result it is convenient to define

$$\bar{\varsigma}_0 \equiv \frac{\alpha_s \left[\bar{\varepsilon} - \varepsilon_L + \theta \int_{\varepsilon_L}^{\varepsilon^n} \left(\varepsilon^n - \varepsilon\right) dG\left(\varepsilon\right)\right]}{\bar{\varepsilon} + \alpha_s \theta \int_{\varepsilon_L}^{\varepsilon^n} \left(\varepsilon^n - \varepsilon\right) dG\left(\varepsilon\right)}$$

$$\hat{\varsigma}_0 \equiv \frac{\alpha_s \int_{\varepsilon^n}^{\varepsilon_H} \left(\varepsilon - \varepsilon^n\right) dG(\varepsilon)}{\bar{\varepsilon} + \alpha_s \int_{\varepsilon_L}^{\varepsilon^n} \left(\varepsilon^n - \varepsilon\right) dG\left(\varepsilon\right)},$$

where $\varepsilon^n \in [\varepsilon_L, \varepsilon_H]$ is the unique solution to

$$G\left(\varepsilon^n\right) = \frac{\Lambda}{1 + \Lambda}.$$

Proposition 11 *Consider the limiting economy (as $\Delta \to 0$) with individual borrowing limit (92), $\alpha_{10} \equiv \alpha_s \alpha$, $\alpha_{11} \equiv \alpha_s (1-\alpha)$, $\alpha_s \in (0,1]$, and $\alpha \in [0,1]$. Let*

$$\varphi_0^n \equiv \lim_{\alpha \to 0} \varphi^n = \bar{\varepsilon} + \alpha_s \theta \int_{\varepsilon_L}^{\varepsilon^n} (\varepsilon^n - \varepsilon)\, dG(\varepsilon).$$

As $\alpha \to 0$,

(i) If $\hat{\varsigma}_0 < \iota < \bar{\varsigma}_0$, then

$$\frac{\mathcal{Z}}{\varphi} \to 0$$
$$\mathcal{V} \to \infty$$
$$\varphi \to \varphi_0^n + \alpha_s (1-\theta) \int_{\varepsilon_L}^{\varepsilon_{10}^*} (\varepsilon_{10}^* - \varepsilon)\, dG(\varepsilon),$$

where $\varepsilon_{10}^* \in (\varepsilon_L, \varepsilon^n)$ is the unique solution to

$$\frac{\theta\left(\varepsilon^n - \varepsilon_{10}^*\right) + (1-\theta)\int_{\varepsilon_{10}^*}^{\varepsilon_H}\left(\varepsilon - \varepsilon_{10}^*\right) dG(\varepsilon) + \theta \int_{\varepsilon^n}^{\varepsilon_H} (\varepsilon - \varepsilon^n) dG(\varepsilon)}{\frac{\bar{\varepsilon}}{\alpha_s} + (1-\theta)\int_{\varepsilon_L}^{\varepsilon_{10}^*}\left(\varepsilon_{10}^* - \varepsilon\right) dG(\varepsilon) + \theta \int_{\varepsilon_L}^{\varepsilon^n} (\varepsilon^n - \varepsilon) dG(\varepsilon)} = \iota.$$

(ii) If $0 < \iota \leq \hat{\varsigma}_0$, then

$$\frac{\mathcal{Z}}{\varphi} \to \frac{G(\varepsilon^*) - [1 - G(\varepsilon^*)]\Lambda}{1 - G(\varepsilon^*)}$$
$$\mathcal{V} \to \frac{\alpha_s G(\varepsilon^*)[1 - G(\varepsilon^*)]}{G(\varepsilon^*) - [1 - G(\varepsilon^*)]\Lambda}$$
$$\varphi \to \bar{\varepsilon} + \alpha_s \int_{\varepsilon_L}^{\varepsilon^*} (\varepsilon^* - \varepsilon)\, dG(\varepsilon),$$

where $\varepsilon^* \in [\varepsilon^n, \varepsilon_H)$ is the unique solution to

$$\frac{\int_{\varepsilon^*}^{\varepsilon_H} (\varepsilon - \varepsilon^*)\, dG(\varepsilon)}{\frac{\bar{\varepsilon}}{\alpha_s} + \int_{\varepsilon_L}^{\varepsilon^*} (\varepsilon^* - \varepsilon)\, dG(\varepsilon)} = \iota.$$

Proposition 11 is analogous to Proposition 8. It considers the limiting economy as the fraction of investors who do not have access to margin loans vanishes, while keeping constant the proportion of investors who can trade equity. In other words, as $\alpha \to 0$, $\alpha_{10} \to 0$ and $\alpha_{11} \to \alpha_s$.

For the limiting economy as $\alpha \to 0$, Figure 4 illustrates in the space of parameters ι (vertical axis) and Λ (horizontal axis), the regions where the equilibria described in parts (i) and (ii) of Proposition 11 exist. Notice that $\bar{\varsigma}_0$ and $\hat{\varsigma}_0$ are functions of ε^n, which is in turn a function of Λ, so to make this dependence explicit, we can write $\bar{\varsigma}_0(\Lambda)$ and $\hat{\varsigma}_0(\Lambda)$. The boundaries $\iota = \bar{\varsigma}_0(\Lambda)$ and $\iota = \hat{\varsigma}_0(\Lambda)$ define three regions.[29] First, if the nominal policy rate is very high, i.e., if $\bar{\varsigma}_0(\Lambda) \le \iota$, then the monetary equilibrium does not exist. Second, if the nominal policy rate is relatively low, i.e., if $0 < \iota \le \hat{\varsigma}_0(\Lambda)$ as in part (*ii*) of Proposition 11, then the aggregate money demand from type 10 investors vanishes in the limit, but the aggregate money demand from type 11 investors with low valuation remains positive in the limit, and therefore real balances and velocity converge to positive limits. Third, if the nominal policy rate is relatively high, i.e., if $\hat{\varsigma}_0(\Lambda) < \iota < \bar{\varsigma}_0(\Lambda)$ as in part (*i*) of Proposition 11, then real balances converge to zero and transaction velocity goes to infinity as $\alpha_{10} \to 0$. The economic rationale for these results is as explained in the context of Proposition 8. And again, the key result is that even though real balances and velocity converge to their nonmonetary equilibrium levels as $\alpha_{10} \to 0$, the real equity price in this cashless limit of the monetary economy exceeds the (corresponding limit of the) nonmonetary-equilibrium price by the value of a resale-option term. Since ε^*_{10} is a function of ι, the asset price is still responsive to monetary policy in the cashless limit, and that the magnitude of this response remains bounded away from zero even though real balances converge to zero.

6 Efficiency and welfare

In this section we pose and solve the planner problems corresponding to the baseline model with fixed capital of Section 2, and the model with capital accumulation of Section 5.1. In both cases we consider a social planner who wishes to maximize the sum of all agents' expected discounted utilities subject to the same meeting frictions that individual agents face in the decentralized formulation. Specifically, in the first subperiod of every period, the planner can only reallocate assets among all equity brokers and the measure $(\alpha_{10} + \alpha_{11}) N_I$ of investors who contact the equity market. We restrict attention to symmetric allocations (identical agents receive equal

[29] It is easy to prove that $\hat{\varsigma}_0(\Lambda) \le \bar{\varsigma}_0(\Lambda)$ for all $\Lambda \ge 0$ (with "=" only if $\Lambda = 0$), and that

$$\lim_{\Lambda\to\infty} \hat{\varsigma}_0(\Lambda) = 0 < \bar{\varsigma}_0(0) = \hat{\varsigma}_0(0) = \frac{\alpha_s(\bar{\varepsilon} - \varepsilon_L)}{\bar{\varepsilon}} \le \lim_{\Lambda\to\infty} \bar{\varsigma}_0(\Lambda) = \frac{\alpha_s[\theta(\varepsilon_H - \varepsilon_L) + (1-\theta)(\bar{\varepsilon} - \varepsilon_L)]}{\bar{\varepsilon} + \alpha_s\theta(\varepsilon_H - \bar{\varepsilon})},$$

where the second inequality is strict unless $\theta = 0$. Also,

$$\frac{\partial \hat{\varsigma}_0}{\partial \varepsilon^n} < 0 \le \frac{\alpha_s \theta G(\varepsilon^n)[\bar{\varepsilon} - \alpha_s(\bar{\varepsilon} - \varepsilon_L)]}{\left[\bar{\varepsilon} + \alpha_s\theta \int_{\varepsilon_L}^{\varepsilon^n} (\varepsilon^n - \varepsilon)\, dG(\varepsilon)\right]^2} = \frac{\partial \bar{\varsigma}_0}{\partial \varepsilon^n}.$$

Hence, $\frac{d\hat{\varsigma}_0}{d\Lambda} < 0 \le \frac{d\bar{\varsigma}_0}{d\Lambda}$.

treatment). For each of the two economies, we also provide a measure of welfare along an equilibrium path, based on the (equally weighted) sum of all agents' expected discounted utilities at the beginning of a period.

6.1 Endowment economy

Let c_t^k and h_t^k denote consumption and production of the homogeneous consumption good in the second subperiod of period t of an agent of type $k \in \{B, E, I\}$. Focus on the environment where only investors are able to carry equity overnight, and let $\tilde{a}_t^I$ denote the beginning-of-period t (before depreciation) equity holding of an individual investor. Let $\bar{a}_t^E$ denote the equity holding of an equity broker at the end of the first subperiod of period t (after OTC trade), and let $\bar{a}_t^I$ denote a measure on $\mathcal{F}([\varepsilon_L, \varepsilon_H])$, the Borel σ-field defined on $[\varepsilon_L, \varepsilon_H]$. The measure $\bar{a}_t^I$ is interpreted as the distribution of post-OTC-trade asset holdings among investors with different valuations who contacted an equity broker in the first subperiod of period t.

With this notation, and letting

$$\Pi \equiv \left\{ \tilde{a}_{t+1}^I, [\bar{a}_t^j, c_t^k, h_t^k]_{j \in \{E,I\}, k \in \{B,E,I\}} \right\}_{t=0}^{\infty},$$

the planner's problem for the model with fixed capital is

$$W^*(y_0) = \max_{\Pi} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \Bigg\{ \Bigg[(\alpha_{10} + \alpha_{11}) \int_{\varepsilon_L}^{\varepsilon_H} \varepsilon y_t \bar{a}_t^I (d\varepsilon) + (1 - \alpha_{10} - \alpha_{11}) \int_{\varepsilon_L}^{\varepsilon_H} \varepsilon y_t a_t^I dG(\varepsilon) \Bigg] N_I + \sum_{k \in \{B,E,I\}} \left(c_t^k - h_t^k \right) N_k \Bigg\} \tag{93}$$

subject to

$$\tilde{a}_t^I N_I \leq A^s \tag{94}$$

$$\bar{a}_t^E N_E + (\alpha_{10} + \alpha_{11}) \int_{\varepsilon_L}^{\varepsilon_H} \bar{a}_t^I (d\varepsilon) N_I \leq (\alpha_{10} + \alpha_{11}) a_t^I N_I \tag{95}$$

$$\sum_{k \in \{B,E,I\}} \left(c_t^k - h_t^k \right) N_k \leq 0 \tag{96}$$

$$a_t^I N_I = \eta \tilde{a}_t^I N_I + (1 - \eta) A^s, \tag{97}$$

and subject to the allocation Π being nonnegative (the expectation operator $\mathbb{E}_0$ is with respect to the probability measure induced by the dividend process). The following proposition characterizes the efficient allocation and the maximum value of the planner's problem.

Proposition 12 *Consider the limiting economy (as $\Delta \to 0$) with exogenous capital. The efficient allocation is characterized by $\bar{a}_t^I(E) = \frac{A^s}{N_I}\mathbb{I}_{\{\varepsilon_H \in E\}}$ for all t, where $\mathbb{I}_{\{\varepsilon_H \in E\}}$ is an indicator function that takes the value 1 if $\varepsilon_H \in E$, and 0 otherwise, for any $E \in \mathcal{F}([\varepsilon_L, \varepsilon_H])$. The welfare achieved by the planner is*

$$\mathcal{W}^*(y_t) = \frac{\psi}{r-g} A^s y_t, \tag{98}$$

where

$$\psi \equiv \bar{\varepsilon} + (\alpha_{10} + \alpha_{11})(\varepsilon_H - \bar{\varepsilon}). \tag{99}$$

According to Proposition 12, the optimal allocation is characterized by the following simple property: among those investors who have a trading opportunity with an equity broker, only those with the highest valuation hold equity shares at the end of the OTC round of trade. Thus the planner conducts the optimal reallocation of first-period post-trade asset holdings given the restrictions of the OTC marketstructure. In this context, ψ can be interpreted as the (flow) shadow value of the asset for the planner, i.e., it is the analogue of φ^n in Proposition 1 or φ in Proposition 2. Recall that, $\varphi^n \le \varphi \le \psi$ (part (i) of Proposition 3).

In the appendix (part (i) of Lemma 27), we show that along the path of a RNE of the limiting continuous-time economy with exogenous capital, welfare is

$$\mathcal{V}^n(y_t) = \frac{\varphi_1^n}{r-g} A^s y_t, \tag{100}$$

where

$$\varphi_1^n \equiv \bar{\varepsilon} + \alpha_{11}\left[\int_{\varepsilon_L}^{\varepsilon^n} (\varepsilon^n - \varepsilon)\, dG(\varepsilon) + \frac{\lambda}{1-\lambda}\int_{\varepsilon^n}^{\varepsilon_H} (\varepsilon - \varepsilon^n)\, dG(\varepsilon)\right], \tag{101}$$

and ε^n satisfies (40). Notice that φ_1^n is the stock price in the RNE of an economy with $\theta = 1$.

In the appendix (part (ii) of Lemma 27), we show that along the path of a RME of the limiting continuous-time economy with exogenous capital, welfare is

$$\mathcal{V}^m(\mathcal{Z}, y_t) = \frac{1}{r-g}\left(u_1^z \frac{\mathcal{Z}}{\varphi} + \varphi_1\right) A^s y_t, \tag{102}$$

where

$$u_1^z \equiv \alpha_{10}\int_{\varepsilon_{10}^*}^{\varepsilon_H} (\varepsilon - \varepsilon_{10}^*)\, dG(\varepsilon) + \alpha_{11}\left[\varepsilon_{11}^* - \varepsilon_{10}^* + \frac{1}{1-\lambda}\int_{\varepsilon_{11}^*}^{\varepsilon_H} (\varepsilon - \varepsilon_{11}^*)\, dG(\varepsilon)\right] \tag{103}$$

$$u_1^s \equiv \alpha_{10}\int_{\varepsilon_L}^{\varepsilon_{10}^*} (\varepsilon_{10}^* - \varepsilon)\, dG(\varepsilon) + \alpha_{11}\left[\int_{\varepsilon_L}^{\varepsilon_{11}^*} (\varepsilon_{11}^* - \varepsilon)\, dG(\varepsilon) + \frac{\lambda}{1-\lambda}\int_{\varepsilon_{11}^*}^{\varepsilon_H} (\varepsilon - \varepsilon_{11}^*)\, dG(\varepsilon)\right], \tag{104}$$

$(\varepsilon_{10}^{*}, \varepsilon_{11}^{*}, \varphi, \mathcal{Z})$ satisfy the equilibrium conditions in Proposition 2, and

$$\varphi_1 \equiv \bar{\varepsilon} + u_1^s = \bar{\varepsilon} + (\alpha_{10} + \alpha_{11})(\varepsilon_{10}^{*} - \bar{\varepsilon}) + u_1^z \tag{105}$$

is the normalized (i.e., multiplied by ρ) stock price in the RME of an economy with $\theta = 1$.

The following result is a corollary of Proposition 2, part (i) of Proposition 3, Proposition 12, Proposition 14, (100), and (102).

Corollary 1 *Consider the limiting economy (as $\Delta \to 0$) with exogenous capital stock. Then*

$$\mathcal{V}^n(y_t) \leq \mathcal{V}^m(\mathcal{Z}, y_t) \leq \mathcal{W}^*(y_t),$$

where the first inequality is strict unless $\iota = \bar{\iota}(\lambda)$, and the second inequality is strict unless $\iota = 0$.

The following result, a corollary of (102)-(104) and Lemma 8, describes welfare in the limiting economy with exogenous capital where all agents have access to credit.

Corollary 2 *Consider the limiting economy (as $\Delta \to 0$) with exogenous capital stock, with $\alpha_{10} \equiv \alpha_s \alpha$, $\alpha_{11} \equiv \alpha_s(1-\alpha)$, $\alpha_s \in (0,1]$, $\alpha \in [0,1]$, and $\lambda \in (0,1]$. As $\alpha \to 0$,*

(i) If $\hat{\varsigma}(0) < \iota < \bar{\varsigma}(0)$, then

$$\lim_{\alpha \to 0} \mathcal{V}^m(\mathcal{Z}, y_t) = \lim_{\alpha \to 0} \mathcal{V}^n(y_t) = \frac{\tilde{\varphi}_1^n}{r-g} A^s y_t.$$

where

$$\tilde{\varphi}_1^n \equiv \lim_{\alpha \to 0} \varphi_1^n = \bar{\varepsilon} + \alpha_s \left[\int_{\varepsilon_L}^{\varepsilon^n} (\varepsilon^n - \varepsilon)\, dG(\varepsilon) + \frac{\lambda}{1-\lambda} \int_{\varepsilon^n}^{\varepsilon_H} (\varepsilon - \varepsilon^n)\, dG(\varepsilon)\right]. \tag{106}$$

(ii) If $0 < \iota \leq \hat{\varsigma}(0)$, then

$$\lim_{\alpha \to 0} \mathcal{V}^n(y_t) < \lim_{\alpha \to 0} \mathcal{V}^m(\mathcal{Z}, y_t) = \frac{\tilde{\varphi}_1^z}{r-g} A^s y_t,$$

where

$$\tilde{\varphi}_1^z \equiv \tilde{\varphi}_1 + \tilde{u}_1^z \frac{G(\varepsilon^*) - \lambda}{1 - G(\varepsilon^*)} \tag{107}$$

$$\tilde{\varphi}_1 \equiv \bar{\varepsilon} + \tilde{u}_1^s \tag{108}$$

with

$$\tilde{u}_1^z \equiv \alpha_s \frac{1}{1-\lambda} \int_{\varepsilon^*}^{\varepsilon_H} (\varepsilon - \varepsilon^*)\, dG(\varepsilon) \tag{109}$$

$$\tilde{u}_1^s \equiv \alpha_s \left[\int_{\varepsilon_L}^{\varepsilon^*} (\varepsilon^* - \varepsilon)\, dG(\varepsilon) + \frac{\lambda}{1-\lambda} \int_{\varepsilon^*}^{\varepsilon_H} (\varepsilon - \varepsilon^*)\, dG(\varepsilon)\right], \tag{110}$$

and ε^* satisfies (74).

Part (i) of Corollary 2 corresponds to the parametrizations characterized in part (i) Proposition 8 for which the limiting economy as $\alpha \to 0$, is cashless. In this case, welfare in the cashless limit of the monetary economy equals welfare in the nonmonetary equilibrium. Thus, although monetary policy affects the stock price in the cashless limit (part (i) of Proposition 8), it does not affect welfare, which is identical to welfare in an economy with no money. This result is in part due to the fact that, since the capital stock, A^s, is exogenous, changes in the market price of capital, ϕ_t^s, have no effect on the allocation of resources in the cashless limit. This result, however, is driven by the fact that the capital stock is exogenous in this formulation.

6.2 Economy with capital accumulation

Next, we turn to the efficient allocation for the economy with capital accumulation. As in Proposition 10, we continue to assume (89). The notation for the planner's problem is as before, except that now we use h_{1t}^I to denote the quantity of general goods used for consumption, h_{2t}^I to denote the quantity of general goods used as input to produce new capital, and $h_t^I = h_{1t}^I + h_{2t}^I$ to denote the labor input (effort) devoted to production of general goods (equal to the quantity of general goods produced). In this case, letting

$$\Pi \equiv \left\{ \tilde{a}_{t+1}^I, c_t^I, h_{1t}^I, h_{2t}^I, \bar{a}_t^I, X_t, \bar{a}_t^E, [c_t^k, h_t^k]_{k \in \{B,E\}} \right\}_{t=0}^{\infty},$$

the planner's problem is

$$W^* (A_0, y_0) = \max_{\Pi} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \Bigg\{ \Bigg[(\alpha_{10} + \alpha_{11}) \int_{\varepsilon_L}^{\varepsilon_H} \varepsilon y_t \bar{a}_t^I (d\varepsilon) + (1 - \alpha_{10} - \alpha_{11}) \int_{\varepsilon_L}^{\varepsilon_H} \varepsilon y_t a_t^I dG(\varepsilon) \Bigg] N_I + \sum_{k \in \{B,E,I\}} \left(c_t^k - h_t^k \right) N_k \Bigg\} \tag{111}$$

subject to

$$\tilde{a}_{t+1}^I N_I \le A_t^s + X_t \tag{112}$$

$$a_{t+1}^I N_I = A_{t+1}^s = \eta (A_t^s + X_t) \tag{113}$$

$$X_t = f_t \left(h_{2t}^I \right) N_I \tag{114}$$

$$\sum_{k \in \{B,E,I\}} c_t^k N_k \le h_{1t}^I N_I + \sum_{k \in \{B,E\}} h_t^k N_k \tag{115}$$

$$\bar{a}_t^E N_E + (\alpha_{10} + \alpha_{11}) \int_{\varepsilon_L}^{\varepsilon_H} \bar{a}_t^I (d\varepsilon) N_I \le (\alpha_{10} + \alpha_{11}) a_t^I N_I, \tag{116}$$

and subject to the allocation Π being nonnegative. Let $\mathcal{X}^*$ denote optimal aggregate investment rate (i.e., $\mathcal{X}^*\Delta$ is optimal investment over a time interval of length Δ).

Proposition 13 *Consider the limiting economy (as $\Delta \to 0$) with capital accumulation. The efficient allocation is characterized by the following conditions: (i) $\bar{a}_t^I(E) = \frac{A_t^s}{N_I}\mathbb{I}_{\{\varepsilon_H \in E\}}$ for all t, where $\mathbb{I}_{\{\varepsilon_H \in E\}}$ is an indicator function that takes the value 1 if $\varepsilon_H \in E$, and 0 otherwise, for any $E \in \mathcal{F}([\varepsilon_L, \varepsilon_H])$; (ii) $\mathcal{X}^* = (\psi/\rho)^{\frac{\sigma}{1-\sigma}} N_I$ for all t, with ψ given by (99); and (iii) the capital stock follows (90) with $\mathcal{X}_t = \mathcal{X}^*$. The welfare achieved by the planner is*

$$\mathcal{W}^*(A_t^s, y_t) = \left[\frac{\psi}{\rho}A_t^s + \frac{1}{r-g}(1-\sigma)\left(\frac{\psi}{\rho}\right)^{\frac{1}{1-\sigma}} N_I\right] y_t. \tag{117}$$

In the setup with capital accumulation the planner optimizes along two margins: the reallocation of the asset, and the investment margin. Optimal reallocation in the OTC trading round is as in the model with exogenous capital, while the optimal investment decision involves equating the marginal rate of substitution between labor and general goods to the optimal shadow price of capital, ψ.

Next, we characterize the welfare function for the economy of Section 5.1). As in Proposition 10, we assume (89).

In the appendix (part (*i*) of Lemma 28), we show that along the path of a RNE of an economy with capital accumulation, welfare is

$$\mathcal{V}^n(A_t^s, y_t) = \left[\frac{\varphi_1^n}{\rho}A_t^s + \frac{1}{r-g}\left(\frac{\varphi_1^n}{\varphi^n} - \sigma\right)\left(\frac{\varphi^n}{\rho}\right)^{\frac{1}{1-\sigma}} N_I\right] y_t, \tag{118}$$

where φ^n is given by (39) (with ε^n given by (40)), φ_1^n is given in (101), and the capital stock follows (90) with $\mathcal{X}_t = (\varphi^n/\rho)^{\frac{\sigma}{1-\sigma}} N_I$.

In the appendix (part (*ii*) of Lemma 28), we show that along the path of a RME of an economy with capital accumulation, welfare is

$$\mathcal{V}^m(\mathcal{Z}, A_t^s, y_t) = \left[\frac{\varphi_1^z}{\rho}A_t^s + \frac{1}{r-g}\left(\frac{\varphi_1^z}{\varphi} - \sigma\right)\left(\frac{\varphi}{\rho}\right)^{\frac{1}{1-\sigma}} N_I\right] y_t, \tag{119}$$

where

$$\varphi_1^z \equiv \varphi_1 + u_1^z \frac{\mathcal{Z}}{\varphi},$$

u_1^z and u_1^s are given by (103) and (104), φ_1 is given in (105), the capital stock follows (90) with $\mathcal{X}_t = (\varphi/\rho)^{\frac{\sigma}{1-\sigma}} N_I$, and $(\varepsilon_{10}^*, \varepsilon_{11}^*, \varphi, \mathcal{Z})$ satisfy the equilibrium conditions in Proposition 2.

The following result is a corollary of Proposition 2, part (i) of Proposition 3, Proposition 13, Proposition 14, (118), and (119).

Corollary 3 *Consider the limiting economy (as $\Delta \to 0$) with capital accumulation and initial condition $(A_t^s, y_t) = (A_0^s, y_0)$. Then*

$$\mathcal{V}^n (A_0^s, y_0) \leq \mathcal{V}^m (\mathcal{Z}, A_0^s, y_0) \leq \mathcal{W}^* (A_0^s, y_0),$$

where the first inequality is strict unless $\iota = \bar{\iota}(\lambda)$ and $\theta = 1$, and the second inequality is strict unless $\iota = 0$.

The following result, a corollary of (119) and Lemma 8, describes welfare in the limiting economy with capital accumulation where all agents have access to credit.

Corollary 4 *Consider the limiting economy with capital accumulation (as $\Delta \to 0$) and initial condition $(A_t^s, y_t) = (A_0^s, y_0)$, with $\alpha_{10} \equiv \alpha_s \alpha$, $\alpha_{11} \equiv \alpha_s (1-\alpha)$, $\alpha_s \in (0,1]$, $\alpha \in [0,1]$, and $\lambda \in (0,1]$. As $\alpha \to 0$,*

$$\lim_{\alpha \to 0} \mathcal{V}^n (A_0^s, y_0) = \left[\frac{\tilde{\varphi}_1^n}{\rho} A_0^s + \frac{1}{r-g} \left(\frac{\tilde{\varphi}_1^n}{\tilde{\varphi}^n} - \sigma \right) \left(\frac{\tilde{\varphi}^n}{\rho} \right)^{\frac{1}{1-\sigma}} N_I \right] y_0,$$

with $\tilde{\varphi}^n$ and $\tilde{\varphi}_1^n$ given by (66) and (106). Moreoever,

(i) If $\hat{\varsigma}(0) < \iota < \bar{\varsigma}(0)$, then

$$\begin{aligned} 0 &\leq \lim_{\alpha \to 0} \left[\mathcal{V}^m (\mathcal{Z}, A_0^s, y_0) - \mathcal{V}^n (A_0^s, y_0) \right] \\ &= \frac{1}{r-g} \left[\left(\frac{\tilde{\varphi}_1^n}{\tilde{\varphi}} - \sigma \right) \left(\frac{\tilde{\varphi}}{\rho} \right)^{\frac{1}{1-\sigma}} - \left(\frac{\tilde{\varphi}_1^n}{\tilde{\varphi}^n} - \sigma \right) \left(\frac{\tilde{\varphi}^n}{\rho} \right)^{\frac{1}{1-\sigma}} \right] N_I y_0, \end{aligned}$$

with "$<$" if $\theta \in [0,1)$, where $\tilde{\varphi}$ given by (69) (with ε^n given by (40), and ε_{10}^* given by (70)).

(ii) If $0 < \iota \leq \hat{\varsigma}(0)$,

$$\begin{aligned} 0 &< \lim_{\alpha \to 0} \left[\mathcal{V}^m (\mathcal{Z}, A_0^s, y_0) - \mathcal{V}^n (A_0^s, y_0) \right] \\ &= \left\{ \frac{1}{\rho} \left(\tilde{\varphi}_1^z - \tilde{\varphi}_1^n \right) A_0^s + \frac{1}{r-g} \left[\left(\frac{\tilde{\varphi}_1^z}{\tilde{\varphi}} - \sigma \right) \left(\frac{\tilde{\varphi}}{\rho} \right)^{\frac{1}{1-\sigma}} - \left(\frac{\tilde{\varphi}_1^n}{\tilde{\varphi}^n} - \sigma \right) \left(\frac{\tilde{\varphi}^n}{\rho} \right)^{\frac{1}{1-\sigma}} \right] N_I \right\} y_0, \end{aligned}$$

where

$$\tilde{\varphi}_1^z \equiv \tilde{u}_1^z \frac{G(\varepsilon^*) - \lambda}{1 - G(\varepsilon^*)} + \tilde{\varphi}_1,$$

with $\tilde{\varphi}$ given by (73) (with ε^* given by (74)), $\tilde{\varphi}_1$ given by (108), and $\tilde{u}_1^z$ given by (109).

In Corollary 4, the thought experiment consists of taking the limit as $\alpha \to 0$ while keeping the capital stock, A_t^s, constant. Part (i) is an economy where money is dominated in rate of return by bonds (i.e., it corresponds to part (i) of Proposition 8). In this case, even keeping A_t^s constant, welfare in the cashless limit of the monetary economy is strictly higher than in the nonmonetary economy, provided $\theta < 1$. In the cashless limit of the monetary economy the equilibrium capital stock is $A_t^s = e^{-\delta t} A_0^s + \left(1 - e^{-\delta t}\right) (\tilde{\varphi}/\rho)^{\frac{\sigma}{1-\sigma}} N_I/\delta$, while in the nonmonetary economy the capital stock is $A_t^s = e^{-\delta t} A_0^s + \left(1 - e^{-\delta t}\right) (\tilde{\varphi}^n/\rho)^{\frac{\sigma}{1-\sigma}} N_I/\delta$, so aggregate consumption (of the dividend good), $C_t = y_t A_t^s$, is higher in the former.

7 Monetary policy, asset prices, and real activity

In this section we study the effects of monetary policy on asset prices and real activity. We first characterize optimal monetary policy, and then turn to positive considerations.

Proposition 14 *As $\iota \to 0$, the recursive monetary equilibrium allocation of the limiting economy (as $\Delta \to 0$), both with exogenous and with endogenous capital stock, converges to the efficient allocation.*

Let $\mathcal{E}_{x|y}$ denote the elasticity of variable x with respect to variable y, i.e., $\mathcal{E}_{xy} \equiv \frac{\partial x}{\partial y}\frac{y}{x}$. The following result characterizes the effect of monetary policy on real asset prices.

Proposition 15 *Consider the limiting economy with capital accumulation (as $\Delta \to 0$). Let $z \equiv \mathcal{Z}/\varphi$, then:*

(i) If $\hat{\iota}(\lambda) < \iota < \bar{\iota}(\lambda)$,

$$\begin{aligned}\mathcal{E}_{\varphi|\iota} &= -\frac{\iota}{\iota + \frac{\alpha_{11}\theta + [\alpha_{10} + \alpha_{11}(1-\theta)]\left[1 - G\left(\varepsilon_{10}^*\right)\right]}{[\alpha_{10}+\alpha_{11}(1-\theta)]G\left(\varepsilon_{10}^*\right)}} \\ &= -\frac{\iota}{\iota + \frac{\alpha_{10} - \alpha_{11}(1-\theta)z}{[\alpha_{10}+\alpha_{11}(1-\theta)]z}},\end{aligned}$$

where ε_{10}^* is given in part (i) of Proposition 2 .

(ii) If $0 < \iota \leq \hat{\iota}(\lambda)$,

$$\begin{aligned}\mathcal{E}_{\varphi|\iota} &= -\frac{\iota}{\iota + \frac{\left[\alpha_{10} + \alpha_{11}\left(1 + \theta\frac{\lambda}{1-\lambda}\right)\right][1 - G(\varepsilon^*)]}{(\alpha_{10}+\alpha_{11})G(\varepsilon^*) - \alpha_{11}\theta\frac{\lambda}{1-\lambda}[1-G(\varepsilon^*)]}} \\ &= -\frac{\iota}{\iota + \frac{\alpha_{10} + \alpha_{11}\left(1+\theta\frac{\lambda}{1-\lambda}\right)}{\alpha_{11}(1-\theta)\frac{\lambda}{1-\lambda} + \left(\alpha_{10} + \alpha_{11}\frac{1}{1-\lambda}\right)z}},\end{aligned}$$

where ε^* is given in part (ii) of Proposition 2.

Proposition 15 provides analytical expressions for the elasticity of the asset price, φ, with respect to the policy rate, ι, both for high and low inflation regimes. In every case the elasticity is negative. In a recursive equilibrium, $\phi_t^m A_t^m = Z A_t^s y_t$, so $z \equiv \mathcal{Z}/\varphi$ as given in Proposition 2 is the value of equilibrium real money balances, $\phi_t^m A_t^m$, relative to the value of the total output, $\phi_t^s A_t^s$, (measured in terms of the dividend good). When written in terms of z, the expressions indicate that keeping the market structure parameters α_{10}, α_{11}, and θ constant, the impact of monetary policy on asset prices would tend to be larger in economies where aggregate real balances are a larger fraction of aggregate output. For example, in the economy with $G(\varepsilon) = G_L \mathbb{I}_{\{\varepsilon_L \leq \varepsilon < \varepsilon_H\}} + \mathbb{I}_{\{\varepsilon_H \leq \varepsilon\}}$ (e.g., from part (i) of Proposition 5),

$$\mathcal{E}_{\varphi|\iota} = -\frac{\iota}{\iota + \frac{\alpha_{11}\theta + [\alpha_{10}+\alpha_{11}(1-\theta)]G_H}{[\alpha_{10}+\alpha_{11}(1-\theta)]G_L}} = -\frac{\iota}{\iota + \frac{\alpha_{10}-\alpha_{11}(1-\theta)z}{[\alpha_{10}+\alpha_{11}(1-\theta)]z}},$$

with $z = \frac{G_L \alpha_{10}}{G_H \alpha_{10} + \alpha_{11}}$.

The following corollary of Proposition 15 reports the elasticity of the real asset price to monetary policy in the limit as $\alpha_{10} \to 0$.

Corollary 5 *Consider the limiting economy (as $\Delta \to 0$) with $\alpha_{10} \equiv \alpha_s \alpha$, $\alpha_{11} \equiv \alpha_s (1-\alpha)$, $\alpha_s \in (0,1]$, $\alpha \in [0,1]$, and $\lambda \in (0,1]$. As $\alpha \to 0$,*

(i) If $\hat{\iota}(\lambda) < \iota < \bar{\iota}(\lambda)$,

$$\mathcal{E}_{\varphi|\iota} \to -\frac{\iota}{\iota + \frac{\theta + (1-\theta)\left[1-G\left(\varepsilon_{10}^*\right)\right]}{(1-\theta)G\left(\varepsilon_{10}^*\right)}}$$

where ε_{10}^* is given in part (i) of Proposition 2 .

(ii) If $0 < \iota \leq \hat{\iota}(\lambda)$,

$$\mathcal{E}_{\varphi|\iota} \to -\frac{\iota}{\iota + \frac{\left(1+\theta\frac{\lambda}{1-\lambda}\right)[1-G(\varepsilon^*)]}{G(\varepsilon^*) - \theta\frac{\lambda}{1-\lambda}[1-G(\varepsilon^*)]}} = \frac{\iota}{\iota + \frac{1-(1-\theta)\lambda}{(1-\theta)\lambda + z}},$$

where ε^* is given in part (ii) of Proposition 2.

The corollary shows that when $\hat{\iota}(\lambda) < \iota < \bar{\iota}(\lambda)$, the elasticity of the asset price with respect to monetary policy is negative and remains bounded away from zero even as z converges to zero. For example, in the economy with $G(\varepsilon) = G_L \mathbb{I}_{\{\varepsilon_L \le \varepsilon < \varepsilon_H\}} + \mathbb{I}_{\{\varepsilon_H \le \varepsilon\}}$ (e.g., from part (i) of Proposition 9),

$$\lim_{\alpha \to 0} \mathcal{E}_{\varphi|\iota} = \begin{cases} -\frac{\iota}{\iota + \frac{\theta + (1-\theta)(1-G_L)}{(1-\theta)G_L}} & \text{if } \hat{\iota}(\lambda) < \iota < \bar{\iota}(\lambda) \\ -\frac{\iota}{\iota + \frac{1-(1-\theta)\lambda}{(1-\theta)\lambda + z}} & \text{if } 0 < \iota \le \hat{\iota}(\lambda), \end{cases}$$

where $z = \frac{G_L - \lambda}{G_H}$ if $0 < \iota \le \hat{\iota}(\lambda)$ (and $z = 0$ if $\hat{\iota}(\lambda) < \iota < \bar{\iota}(\lambda)$). In this case, for instance, if $\hat{\iota}(\lambda) < \iota < \bar{\iota}(\lambda)$, then $\mathcal{E}_{\varphi|\iota} \to 0$ if $\theta G_L = 0$, and $\mathcal{E}_{\varphi|\iota} \to -1$ if $(1-\theta) G_L = 1$.

To conclude, notice that in the economy with capital accumulation with production technology given by (89), the elasticity of investment with respect to ι is

$$\mathcal{E}_{\mathcal{X}|\iota} = \frac{\sigma}{1-\sigma} \mathcal{E}_{\varphi|\iota}.$$

8 Quantitative analysis

We regard a unit of time as corresponding to one day. The number of outstanding shares is normalized to 1, i.e., $A^s = 1$. The dividend growth rate is independently lognormally distributed over time, with mean .04 and standard deviation .12 per annum (e.g., as documented in Lettau and Ludvigson (2005), Table 1). That is, $y_{t+1} = e^{x_{t+1}} y_t$, with $x_{t+1} \sim \mathcal{N}(g, \Sigma^2)$, where $g = \mathbb{E}(\log y_{t+1} - \log y_t) = .04$ and $\Sigma = SD(\log y_{t+1} - \log y_t) = .12$. Over the sample period 1994-2007, the average nominal policy rate was 4.47% per annum, and the average inflation rate was 2.69% per annum.[30] Thus, we set $\rho^p = .0447$ and $\bar{\pi} \equiv \pi - g = .0269$, implying a real rate of $r = \rho^p - \bar{\pi} = .0178$ per annum.[31] The parameter δ can be taken as a proxy of the riskiness of stocks; we choose $\delta = .3$, i.e., a productive unit has a 70 percent probability of remaining productive each year. The distribution of idiosyncratic valuations, G, is assumed to be lognormal with parameters chosen so that, given the rest of the parametrization, the asset price falls by about 11 basis points in response to a 1 basis point increase in the nominal policy rate in the economy with no credit (i.e., if either $\lambda = 0$ or $\alpha_{11} = 0$). In other words,

[30] For the policy rate we use the 3-month Eurodollar futures rate (series IEDCS00 produced by the CME Group available via Datastream). The annual average inflation rate is imputed as $[CPI(January\_2008)/CPI(January\_1994)]^{1/14} - 1$, where $CPI(Month\_Year)$ is monthly CPI index available from FRED at https://fred.stlouisfed.org/series/CPIAUCSL.

[31] To streamline the presentation, here we assume r is constant and therefore associate changes in the policy rate ρ^p with changes in π. In Lagos and Zhang (2017), which corresponds to a special case of this model with no credit, i.e., if either $\lambda = 0$ or $\alpha = 1$, we allow for the possibility that when the policy rate changes by $\Delta\rho^p$, the real rate changes by $\Delta r = w\Delta\rho^p$ and the inflation rate changes by $\Delta\pi = (1-w)\Delta\rho^p$, where $w \in [0,1]$ indexes the degree of passthrough from nominal rates to real rates.

the discipline in our calibration strategy for G and δ consists of ensuring that the no-credit economy generates asset price responses to money shocks that match the empirical estimates in Lagos and Zhang (2017). We adopt the formulation with $\alpha_{10} \equiv \alpha_s \alpha$ and $\alpha_{11} \equiv \alpha_s (1 - \alpha)$, and set $\alpha_s = 1$, so every agent is able to trade equity every day. Our quantitative exercises consist of reporting asset price responses to changes in ρ^p for all $(\alpha, \lambda, \theta) \in [0,1]^3$. Specifically, we focus on the semi-elasticity of the asset price, ϕ^s, with respect to the policy rate, ρ^p. Since the value is always negative, we report $\mathcal{S} = \left| \frac{d\phi^s/\phi^s}{d\rho^p} \right|$. When an exercise calls for a baseline value of λ or θ, we use .5.

Figure 5 reports $\mathcal{S}$ for economies indexed by $(\alpha, \lambda) \in [0,1] \times \{.50, .75, .90, .99\}$. Our baseline calibration ensures that, in the no-credit economy with $\alpha = 1$, $\mathcal{S} \approx 11$ for all λ. The main finding is that the response of the asset price to nominal rate shocks remains significant even in the limiting economy as $\alpha \to 0$. For example, the semi-elasticity is larger than 5 as $\alpha \to 0$ even if investors are able to leverage up to 100 times their wealth. Figure 6 reports $\mathcal{S}$ for economies indexed by $(\alpha, \theta) \in [0,1] \times \{.10, .50, .75, .99\}$. Again, the response of the asset price to nominal rate shocks remains significant in the limiting economy as $\alpha \to 0$. For example, the semi-elasticity is about 9.75 as $\alpha \to 0$ even if investors are able to capture 99% of the gains from trade in trades intermediated by bond brokers. Figure 7 reports $\mathcal{S}$ for economies indexed by $(\alpha, \rho^p) \in [0,1] \times \{.02, .03, .04, .06\}$. This exercise shows that for every level of α, the asset price response tends to be smaller in environments with a higher background nominal rate. Recall that $\rho^p = .04$ is our baseline (based on the average nominal rate between 1994 and 2007). The response of the asset price to nominal rate shocks remains significant in the limiting economy as $\alpha \to 0$, even in the context of an economy with a nominal rate that is significantly higher than the baseline. For example, $\mathcal{S}$ is about 7 in the limiting economy as $\alpha \to 0$ when $\rho^p = .06$ (a policy rate fifty percent higher than the historical average).

Figures 8, 9, and 10 offer a comprehensive picture of the magnitude of the effects of monetary policy in limiting economies with $\alpha \to 0$. For a wide range of economies indexed by a pair ρ^p and $\lambda \in [0,1]$, Figure 8 reports the value of $\mathcal{S}$ in the limit of the monetary equilibrium as $\alpha \to 0$. The level sets in the right panel show it is not easy to find reasonable parametrizations that imply a value of $\mathcal{S}$ below 4. Figures 9 and 10 tell a similar story.

9 Discussion

9.1 On the moneyless approach to monetary economics

Our results on the medium-of-exchange role of money in the transmission of monteary policy run counter to a large body of work that follows a moneyless approach to monetary economics. This moneyless approach was advocated by Woodford (1998) and, based on the treatments in Woodford (2003) and Gali (2008), is now considered by many "the textbook" approach to monetary theory and practice. The common justification for doing monetary economics without money is the view that the frictions associated with the medium-of-exchange role are irrelevant in the transmission of monetary policy. This sweeping view rests on two specific results. Both results rely on a model where the medium-of-exchange role of money is not explicit, but rather is proxied by either assuming money is an argument of a utility function, or by imposing that certain purchases be paid for with cash acquired in advance. The first result is theoretical, and can be found in Woodford (1998). The second result is quantitative, and can be found in Woodford (2003) and Gali (2008). We discuss each of these results in turn.

Woodford (1998) considers a version of the cash-in-advance economy of Lucas (1980) with "cash goods" and "credit goods" as in Lucas and Stokey (1983), but where the set of cash goods is represented with a parameter $\alpha \in (0, \gamma]$ for some $\gamma \in (0, 1)$. The economy with $\alpha \to \gamma$ corresponds to the formulation with no credit goods of Lucas (1980). When $\alpha \to 0$, the economy is interpreted to be approaching a "cashless limit" where there are no cash goods, i.e., a conventional perfectly competitive nonmonetary model without a cash-in-advance constraint. In this context, the first result in Woodford (1998) is that under the assumption that the money supply sequence $\{M_t\}_{t=0}^{\infty}$ satisfies $M_t \geq \underline{M}$ for some $\underline{M} > 0$ for all t, then there is no monetary equilibrium in the limiting case $\alpha = 0$ (in the sense that the nominal price of cash goods, $\{p_t\}_{t=0}^{\infty}$, diverges to infinity). The second result is that given p_t is finite for all $\alpha < 0$, one cannot find a solution for the limiting case $\alpha = 0$ as an approximation to the small-α case. Woodford interprets this result to mean that in this model "the use of money in transactions is intrinsic to the model's ability to determine an equilibrium price level." Woodford then augments the model by assuming the government adopts a fiscal-monetary regime that ensures money is valued and held by private agents even if it is merely a redundant asset. Specifically, the government is assumed to: (i) maintain a strictly positive level of *nominal* government liabilities (so that *cash* taxes must be levied on the private sector in order to service the nominal debt), and (ii) pay a nominal interest on money balances (equal to the nominal interest on the government debt), where the nominal interest rate follows an exogenous rule described by a function of $g(\cdot)$ of p_t, assumed to be continuously differentiable in the neighborhood of some p^*, with $g(p^*)$ chosen to ensure that money is held in the equilibrium of the economy with $\alpha = 0$ (i.e., to ensure the Euler equation for money holds with equality, and the relevant transversality condition satisfied given $p_{t+1} = p_t = p^*$ for all t). Condition (ii) effectively makes money and bonds the same asset (with the same rate of return), which ensures private agents are willing to hold money even though it is not useful in transactions. Notice that since money plays no role as a medium of exchange, there is no demand of money for private transactions that can be equated to the money supply to determine p_t. Condition (i), however, amounts to assuming a private-sector demand for money needed to meet the nominal tax liabilities with the government; this tax-induced money demand allows the price level, p_t, to be determined using the government budget constraint. In the context of the cash-credit cash-in-advance model under the fiscal-monetary regime described by conditions (i) and (ii), Woodford shows the central approximation result of his paper, namely that the equilibrium is continuous in the parameter α, i.e., the equilibrium of the economy with $\alpha = 0$ can be well approximated by the equilibrium of an economy with positive but small enough value of α.

Hence, against the background of a cash-in-advance economy subject to assumptions (i) and (ii), the cashless limit just described, i.e., the economy with $\alpha = 0$ where money is a redundant asset with no role in exchange, can be regarded as a good approximation to a monetary economy where money is needed to satisfy a cash-in-advance constraint but only for a very small set of goods, i.e., the economy with α is postitive but very small. Since there are no monetary variables in the Euler equations of the limiting economy, this approximation result is used to justify neglecting monetary variables in Euler equations more generally, alluding to economies with "highly developed financial institutions" (meaning economies with low α). In this context, for the Euler equations for other durable assets, ignoring monetary variables is equivalent to simply assuming a period utility function of the form $U(c, m) = u(c) + Av(m)$ of consumption, c, and real money balances, m, for given functions $u(\cdot)$ and $v(\cdot)$, and a constant $A \in \mathbb{R}_+$. Thus, the Woodford cashless-limit approximation result is often used to justify this specific money-in-the-utility-function formulation, sometimes with $A \approx 0$. In sum, the takeaway of Woodford (1998) is that the cashless equilibrium in the limiting case, which is independent of money demand for transactions, can be used to approximate the monetary equilibrium in any case in which medium-of-exchange frictions exist but are small.

The textbook treatments of monetary policy in Woodford (2003) and Gali (2008) assign a very limited role to money. For the most part, the medium-of-exchange role is either ignored, or when it is acknowledged, it is incorporated implicitly by assuming real money balances as an argument of the agents' utility functions (or some equivalent cash-in-advance formulation). The preferred specification is $U(c, m) = u(c) + Av(m)$. This separable specification is justified by showing that, in the context of a competitive model with no credit frictions, if $U(c, m)$ is nonseparable, then the elasticity of output with respect to a monetary shock that raises the nominal interest rate by one percentage point is proportional to inverse velocity, $M_t / (p_t Y_t)$, where M_t / p_t denotes aggregate real money balances, and Y_t denotes GDP. Woodford (2003, p. 113) and Gali (2008, p. 31) argue that since $M_t / (p_t Y_t)$ is small in the data (e.g., with M_t interpreted as the monetary base), the effect of monetary policy on output that is attributable to monetary frictions is quantitatively small so it can be ignored, e.g., by considering the simpler formulation $U(c, m) = u(c) + Av(m)$, often even assuming $A \approx 0$.

The literature mentions several reasons why it may be interesting to study monetary policy in limit cashless economies such as the one with $\alpha \to 0$ in Woodford (1998). The first, as argued by Woodford (1998, p. 174), is that the hypothetical cashless limit may one day become a reality as a result financial innovations that continually reduce the quantity of the monetary base that needs to be held on average to carry out a given volume of transactions: "The only natural limit to this process is an ideal state of frictionless financial markets in which there is no positive demand for the monetary base at all, if it is dominated by other financial assets, and no determinate demand for it if it is not." The second, is that the cashless limit may be a useful thought experiment, as argued by Wicksell (1898, p. 70) when considering his "pure credit economy," defined as:

> *"... a state of affairs in which money does not actually circulate at all, neither in the form of coin (except perhaps as small change) nor in the form of notes, but where all domestic payments are effected by means of the Giro system and bookkeeping transfers. A thorough analysis of this purely imaginary case seems to me to be worth while, for it provides a precise antithesis to the equally imaginary case of a pure cash system, in which credit plays no part whatsoever. The monetary systems actually employed in various countries can then be regarded as combinations of these two extreme types. If we can obtain a clear picture of the causes responsible for the*

value of money in both of these imaginary cases, we shall, I think, have found the right key to a solution of the complications which monetary phenomena exhibit in practice."

The cashless limit we considered (e.g., Proposition 8) is in the spirit of Wicksell's "pure credit economy" and in line with the motivation for Woodford's cashless limit. Generically, however, our results stand in contrast with Woodford's: we find that in general the medium-of-exchange role of money is important for monetary transmission, and remains a significant conduit for monetary policy even in the cashless limit. As $\alpha \to 0$, real balances converge to zero, transaction velocity goes to infinity, and the monetary economy converges to a limit where monetary policy still has significant effects on welfare, asset prices, consumption, investment, and output. There is one special case of our theory that delivers irrelevance results for the medium-of-exchange role of money in the cashless limit that are similar to Woodford's. It is the case where financial intermediaries have no market power, i.e., $\theta = 1$. So, in order to argue that monetary frictions are irrelevant in cashless limiting economies or almost irrelevant near-cashless economies, it is necessary to also adopt the view that investors are always able to reap the entire share of the gains from trade when interacting with intermediaries in financial markets.[32] Our theoretical point here is that $\theta = 1$ is nongeneric. Whether the perfectly competitive case with $\theta = 1$ is the relevant case for applied work, is likely to be ultimately an empirical issue that deserves further study. We are aware of no evidence that $\theta = 1$ is the norm empirically, even in the financially advanced economies with low levels inverse velocity of the monetary base that Woodford (2003) and Gali (2008) argue are well approximated by the moneyless approach to monetary policy.[33]

For $\theta < 1$, our theory provides counterexamples to the typical claims (based on the reduced-form arguments outlined above) commonly used to endorse the moneyless approach. For exam-

[32] Our formulation also assumes individual investors are subject to borrowing limits; a collateral constraint in the model of Section 2, and a given borrowing limit in the model of Section 5.2). This can be interpreted as a second departure from what may be described as "frictionless financial markets." However, our cashless results hold even if $\lambda \to 1$ in the economy of Section 2, and if $\Lambda \to \infty$ in the economy of Section 5.2.

[33] Notice that for a given $\theta < 1$, our cashless limiting economy is different depending on the underlying monetary policy, ι, and credit conditions, as captured by λ in the baseline model. If the policy rate and leverage are relatively large as in part (i) of Proposition 8 then money is dominated in rate of return, real balances converge to zero, and asset prices respond to ι even in the cashless limit. If the policy rate and leverage are relatively low as in part (ii) of Proposition 8, then as $\alpha \to 0$, real balances are not dominated in rate of return, do not converge to zero, and monetary policy remains effective in the limit. In both cases our conclusions regarding the effects of ι on prices and allocations in the pure-credit limiting economy is clearly at odds with the limiting irrelevance result in Woodford (2003).

ple, Woodford (2003, p. 32) claims that the basic model in his book "abstracts from monetary frictions, in order to focus attention on more essential aspects of the monetary transmission mechanism...". Gali (2008, p. 10) claims that "...there is generally no need to specify a money demand function, unless monetary policy itself is specified in terms of a monetary aggregate, in which case a simple log-linear money demand shcedule is postulated." The results we presented above (in particular those in Section 7 and Section 8) and in related work (Lagos and Zhang, 2017) indicate that traditional medium-of-exchange considerations are in fact an essential aspect of the monetary transmission mechanism—even in the cashless limit or in near-cashless economies in which liquidity-saving mechanisms have developed sufficiently to make the inverse velocity of the monetary base very small. Any attempt to assess the macroeconomic effects of monetary policy without such considerations is necessarily incomplete.

9.2 On reduced-form models of money demand

There are well known critiques of reduced-form models of money. Kareken and Wallace (1980) for example, state two. The first, is that assuming money is an argument of a utility (or a production) function is an instance of "implicity theorizing," by which they mean that while there may be stories that can be told to justify the approach (e.g., that money provides unmodelled transaction services), the assumptions implicit in these stories cannot be regarded as primitives, and unless the underlying environment is made explicit, the internal consistency of the theory cannot be assessed. The second criticism is that reduced-form specifications beg too many questions, explain too little. What is the thing called "money"? Is it a private liability? A government liability? A commodity? If it is a government liability, which one? If there are many countries, does the liability issued by the government of country A enter the utility function of a citizen of country B? Compelling as they may be, these two criticisms are ignored in most of applied monetary economics. The reason, we suspect, is that these criticisms may not seem too serious in practice. Suppose one wants to study the effects of a monetary policy shift in an advanced economy like the United States. The common practitioner's view would be that in this context, there are some reasonable choices for the assets that play the role of money, and that the unmodelled medium-of-exchange frictions subject to the Kareken-Wallace implicit theorizing critique are likely to be small anyways.[34]

[34]The arguments of the reduced-form utility functions used in practice typically include a meaasure of real money balances (e.g., as in Sidrauski (1967), Gali (2008), and Woodford (2003)), but also government bonds, equity shares, and other financial assets (e.g., as in Krishnamurthy and Vissing-Jorgensen (2012)).

The near-cashless results in Woodford (2003) and Gali (2008) differ from ours because they rely on a money-in-the-utility (MIU) formulation that can sometimes fail to capture important aspects of the role money plays in transactions. For this reason it may be useful to show the limitations of their reduced-form approach in the context of our economic environment. We want to stress that these limitations are relevant even after we have agreed on which assets should be included in the utility function, and remain relevant even for practical, routine questions in monetary policy in the context of advanced economies with highly developed liquidity-saving mechanisms and credit-based payment arrangements resulting in low aggregate real balances.

The equilibrium conditions for our model can be obtained from the following reduced form:

$$\max_{\{c_t,h_t,\tilde{\boldsymbol{a}}_{t+1}\}} \mathbb{E}\sum_{t=0}^{\infty}\beta^t\left[\mathcal{U}\left(c_1,c_2;\boldsymbol{s}_t\right)+c_t-h_t\right] \tag{120}$$

$$\begin{aligned}\text{s.t. } c_t+\phi_t^s\tilde{a}_{t+1}^s+\phi_t^m\tilde{a}_{t+1}^m &= h_t+\left(\bar{\varepsilon}y_t+\phi_t^s\right)a_t^s+\phi_t^m a_t^m+T_t\\ c_{1t} &= \frac{a_t^m}{p_t}y_t\\ c_{2t} &= a_t^s y_t\\ \boldsymbol{a}_{t+1} &= \left(\tilde{a}_{t+1}^m,\eta\tilde{a}_{t+1}^s+(1-\eta)A^s\right),\end{aligned}$$

with

$$\mathcal{U}\left(c_1,c_2;\boldsymbol{s}_t\right)\equiv\overline{\mathcal{U}}_{1t}c_1+\overline{\mathcal{U}}_{2t}c_2,$$

where $\boldsymbol{s}_t\equiv\left(y_t,p_t,q_t,\boldsymbol{\phi}_t,\varepsilon_{10t}^*,\varepsilon_{11t}^*\right)$, $\overline{\mathcal{U}}_{1t}\equiv\frac{p_t}{y_t}\left(\bar{v}_{It}^m-\phi_t^m\right)$, $\overline{\mathcal{U}}_{2t}\equiv\frac{1}{y_t}\left[\bar{v}_{It}^s-\left(\bar{\varepsilon}y_t+\phi_t^s\right)\right]$, and $\bar{v}_{It}^k\equiv\int v_{It}^k\left(\varepsilon\right)dG(\varepsilon)$ for $k\in\{m,s\}$ is defined in Lemma 5.

If we focus on recursive equilibrium, then $\mathcal{U}\left(c_1,c_2;\boldsymbol{s}_t\right)=U\left(c_1,c_2;\boldsymbol{x}_t\right)$, where

$$U\left(c_1,c_2;\boldsymbol{x}\right)\equiv u^z c_1+u^s c_2, \tag{121}$$

where $\boldsymbol{x}\equiv\left(\phi^s,\varepsilon_{10}^*,\varepsilon_{11}^*\right)$, and u^z and u^s are given by

$$\begin{aligned}u^z &= \left[\alpha_{10}+\alpha_{11}\left(1-\theta\right)\right]\int_{\varepsilon_{10}^*}^{\varepsilon_H}\left(\varepsilon-\varepsilon_{10}^*\right)dG(\varepsilon)\\ &\quad+\alpha_{11}\theta\left[\varepsilon_{11}^*-\varepsilon_{10}^*+\frac{1}{1-\lambda}\int_{\varepsilon_{11}^*}^{\varepsilon_H}\left(\varepsilon-\varepsilon_{11}^*\right)dG(\varepsilon)\right]\end{aligned} \tag{122}$$

$$\begin{aligned}u^s &= \left[\alpha_{10}+\alpha_{11}\left(1-\theta\right)\right]\int_{\varepsilon_L}^{\varepsilon_{10}^*}\left(\varepsilon_{10}^*-\varepsilon\right)dG\left(\varepsilon\right)\\ &\quad+\alpha_{11}\theta\left[\int_{\varepsilon_L}^{\varepsilon_{11}^*}\left(\varepsilon_{11}^*-\varepsilon\right)dG\left(\varepsilon\right)+\frac{\lambda}{1-\lambda}\int_{\varepsilon_{11}^*}^{\varepsilon_H}\left(\varepsilon-\varepsilon_{11}^*\right)dG\left(\varepsilon\right)\right].\end{aligned} \tag{123}$$

Notice that, u^z and u^s are implicit functions of $\boldsymbol{x} \equiv (\phi^s, \varepsilon^*_{10}, \varepsilon^*_{11})$.[35] The first-order conditions for (120) are

$$\phi^m_t \geq \beta \mathbb{E}_t \left[\frac{\partial U(c_{1t+1}, c_{2t+1}; \boldsymbol{x}_{t+1})}{\partial c_{1t+1}} \frac{y_{t+1}}{p_{t+1}} + \phi^m_{t+1} \right], \text{ with `` = '' if } 0 < \tilde{a}^m_{t+1}$$
$$\phi^s_t = \beta \eta \mathbb{E}_t \left[\left(\bar{\varepsilon} + \frac{\partial U(c_{1t+1}, c_{2t+1}; \boldsymbol{x}_{t+1})}{\partial c_{2t+1}} \right) y_{t+1} + \phi^s_{t+1} \right].$$

If we focus on a recursive equilibrium of the discrete-time economy with period length Δ, and let $\Delta \to 0$, these first-order conditions become

$$\iota \geq \frac{u^z}{\varphi}, \text{ with `` = '' if } 0 < \mathcal{Z} \tag{124}$$
$$\varphi = \bar{\varepsilon} + u^s. \tag{125}$$

If money is held, (124) and (125) imply

$$\frac{u^z}{\bar{\varepsilon} + u^s} = \iota.$$

The right side of this condition is the opportunity cost of carrying a dollar between periods, and the left side is the marginal rate of substitution between money and equity.

If we substitute (122) and (123), (124) and (125) become identical to the last two conditions in Lemma 13, so formulation (120) is in this sense equivalent to our micro model of investor behavior. Hence, we can define recursive equilibrium as before: a vector, $(\varepsilon^*_{10}, \varepsilon^*_{11}, \varphi, \mathcal{Z})$, that satisfies the first two conditions in Lemma 13, together with (124) and (125) (taking into account that u^z and u^s are given by (122) and (123)).

Suppose instead, that we regard u^z and u^s as a pair of fixed parameters, as would be the case in a prototypical MIU formulation that ignores all the micro details of the OTC trading round, and instead proxies for them with a utility function U given by (121) that is taken parametrically and is meant to capture the "convenience yield" or "liquidity services" of certain asssets. One obvious problem with this approach is that, instead of regarding U as an *indirect utility function* (i.e., and equilibrium object) it treats U as if it were a primitive, i.e., exogenous and invariant to changes in the policy and marketstructure parameters. In contrast, in a structural micro-founded model such as the one we develop in Section 2, changes in the policy rate ι, credit conditions λ, and marketstructure parameters $(\alpha_{ks})_{k,s\in\{0,1\}}$ and θ, change

[35] Instead of using $\boldsymbol{x}$ to index U, we could use the set of all parameters, $\Pi = (\beta, \eta, \bar{\gamma}, G, \theta, (\alpha_{ks})_{k,s\in\{0,1\}}, \lambda, \mu)$ since the equilibrium objects $(\phi^s, \varepsilon^*_{10}, \varepsilon^*_{11})$ are themselves constant functions of Π in a recursive equilibrium.

the shape of the indirect utility function U through their effects on $\boldsymbol{x}$, or more explicitly, through their effects on u^z and u^s, that are functions of those parameters, as is clear from (122) and (123). An equilibrium for the MIU economy would be a pair, $(\varphi, \mathcal{Z})$, that satisfies (124) and (125) taking u^z and u^s parametrically. For this MIU economy, the equilibrium is simple: The equity price is given explicitly by (125); it consists of the expected dividend, $\bar{\varepsilon}$, and the "convenience yield" of equity, u^s. A robust feature of this formulation (exploited in the work that advocates studying monetary policy ignoring money-demand considerations) is that because U is separable in real balances, the equilibrium level of real balances plays no role in the determination of the rest of the equilibrium. In this case, for instance, the real asset price, φ, is independent of $\mathcal{Z}$ and therefore, independent of the policy rate, ι. Equilibrium real balances are given by (124), namely, monetary equilibrium exists only if $\iota = u^z/\varphi$, and in this case the equilibrium aggregate real balance is any $\mathcal{Z} \in [0, \infty)$.[36] The upshot is that the properties of the equilibrium prices $(\varphi, \mathcal{Z})$ that result from treating U as a primitive are in general very different from the properties of the equilibrium prices of the model where the role of money and equity in the exchange process is accounted for explicitly.

Obvious though this "Lucas critique" type of observation may be, it also turns out to be a critical shortcoming of the reduced-form approach, especially when used to draw conclusions on the importance of the medium-of-exchange function of money and its role in the transmission of monetary policy. To give a concrete example, based on the MIU formulation one would conclude that φ is independent of monetary policy. However, once the exchange process through which money yields liquidity services is spelled out, e.g., as in the model of Section 2, one learns that φ is decreasing in ι because u^s is really a decreasing function of ι. As another example, consider a cashless limit or a near-cashless economy like the ones often used to justify ignoring monetary frictions in the New Keynesian textbooks. The way such a cashless limit (e.g., resulting from a sequence of improvements in liquidity-saving mechanisms or credit-based payment arrangements) can be captured in the reduced-form model is by driving the marginal value of real balances to zero in the assumed utility function. For example, in our case with U given by (121), suppose we start with a policy rate consistent with monetary equilibrium, i.e., $\iota = u^z/\varphi$. Then as the parameter u^z is reduced, real balances go to zero but this has no effect on the rest of the equilibrium; in particular the real equity price (125) is unchanged because u^s

[36]The indeterminacy of $\mathcal{Z}$ in the monetary equilibrium in this example is due to the fact that U is not only separable, but also linear in real balances. If U where separable but not linear in real balances, then (125) would be unchanged and (124) (at equality) would determine a unique $\mathcal{Z}$.

is taken to be a parameter in the MIU formulation. The irrelevance of the medium-of-exchange role of money is as strong as it can be in this MIU formulation: the rest of the equilibrium (i.e., asset prices, and in the model with capital accumulation, also consumption, investment, and output) are invariant to $\mathcal{Z}$ in the cashless limit, but also away from it. In the micro-founded model, in contrast, u^s is a function of the credit conditions λ, the marketstructure parameters $(\alpha_{ks})_{k,s\in\{0,1\}}$ and θ, and in particular a function of the policy rate ι—even in the cashless limit, e.g., as $\mathcal{Z} \to 0$ because $\alpha \to 0$ as in part (i) of Proposition 8.

In sum, the widespread New Keynesian view that medium-of-exchange monetary frictions are unimportant for monetary transmission relies on two irrelevance results that are based on: (i) a reduced-form specification of cash and credit transactions that fails to capture the effects of monetary policy on prices and allocations that remain significant even in near-cashless economies, and (ii) a presumption that the financial markets implicit in the reduced-form specification are frictionless (and in particular that investors are always able to reap the entire share of the gains from trade when trading with intermediaries in those unmodelled cash and credit markets).

We have shown that if the role of money as a medium of exchange is modeled explicitly and financial markets exhibit realistic frictions (e.g., credit takes the form of collateralized loans intermediated by brokers who have at least some degree of market power), then monetary policy conducted by means of changing a nominal interest rate is transmitted to the real allocations—even in the cashless limit, i.e., even if access to credit is so generalized that real balances are negligible. Along this cashless limit, the path of the monetary aggregate (and real balances) become irrelevant but the monetary frictions that give money value as a medium of exchange remain important for the transmission of monetary policy. These monetary frictions are resilient in that they remain operative even when real balances are negligible as is the case in the cashless limit. The reasons are: (i) the true indirect utility function for money and consumption is changing along the cashless limit, so MIU formulations that assume the utility function is stable to policy or changes in the market structure are not be able represent the limit of the monetary equilibrium of a model such as ours, where money and intermediated credit help agents overcome commitment and double-coincidence-of-wants frictions. (ii) The financial market in our model is not frictionless: credit is limited and involves intermediaries who have some degree of market power.

9.3 Price level determination in the cashless limit

The price level determination in limiting cashless economies is a recurrent theme in Woodford (2003), so for comparison purposes, it may be useful to explain the behavior of the price level along the cashless limit of our economy. In a RME, the price level (measured by the nominal price of equity shares) is[37]

$$p_t = \frac{\varepsilon_{10}^* + \phi^s}{ZA^s} A_t^m.$$

In the discrete-time formulation with period length is Δ, $p_t = [\varepsilon_{10}^* + \Phi^s(\Delta)] A_t^m / [\mathrm{Z}(\Delta) A^s]$, so as $\Delta \to 0$, we get

$$p_t = \frac{\varphi}{\mathcal{Z}A^s} A_t^m = \begin{cases} \frac{[1-G(\varepsilon_{10}^*)]\alpha_{10}+\alpha_{11}}{\alpha_{10}G(\varepsilon_{10}^*)} \frac{A_t^m}{A^s} & \text{if } \hat{\iota}(\lambda) < \iota < \bar{\iota}(\lambda) \\ \frac{[1-G(\varepsilon^*)]\left(\alpha_{10}+\alpha_{11}\frac{1}{1-\lambda}\right)}{\alpha_{10}G(\varepsilon^*)+\alpha_{11}\frac{1}{1-\lambda}[G(\varepsilon^*)-\lambda]} \frac{A_t^m}{A^s} & \text{if } 0 < \iota \le \hat{\iota}(\lambda). \end{cases}$$

Then if we let $\alpha_{10} \equiv \alpha_s \alpha$ and $\alpha_{11} \equiv \alpha_s (1-\alpha)$,

$$\lim_{\alpha \to 0} p_t = \begin{cases} \frac{1}{A^s G(\varepsilon_{10}^*)} \lim_{\alpha \to 0} \frac{A_t^m}{\alpha} & \text{if } \hat{\iota}(\lambda) < \iota < \bar{\iota}(\lambda) \\ \frac{1-G(\varepsilon^*)}{G(\varepsilon^*)-\lambda} \frac{A_t^m}{A^s} & \text{if } 0 < \iota \le \hat{\iota}(\lambda). \end{cases}$$

If $0 < \iota \le \hat{\iota}(\lambda)$, then $\lim_{\alpha \to 0} p_t$ is necessarily finite for any path $\{A_t^m\}_{t=0}^\infty$. If the monetary authority wishes to implement a certain price path for a monetary equilibrium of the cashless limiting economy with $\iota \in (\hat{\iota}(\lambda), \bar{\iota}(\lambda))$, then it can simply choose a money supply process $\{A_t^m\}_{t=0}^\infty$ given by

$$A_t^m = \alpha M_t \text{ with } \dot{M}_t = \pi M_t,$$

which implements a price level in the cashless limit that is equal to

$$\lim_{\alpha \to 0} p_t = \frac{M_t}{A^s G(\varepsilon_{10}^*)}.$$

By choosing the level of M_0, the monetary authority can implement any price level in the RME of the cashless limiting economy. Intuitively, the monetary authority can always implement a price level that remains well defined (i.e., finite) even in the cashless limit, simply by keeping the money supply *per investor of type* 10 stable along the cashless limit.

[37]The price level measured by the nominal price of general goods is just $\frac{1}{\phi_t^m} = \frac{p_t}{(\varepsilon_{10}^* + \phi^s) y_t}$, so we focus the analysis on p_t.

10 Conclusion

We conclude by mentioning what we think are three promising avenues for future work. First, the model we have presented is tailored to the transmission of monetary policy that operates through financial markets. We have chosen this route guided by the empirical and quantitative results in Lagos and Zhang (2017). Here we have shown how, through its effect on asset prices, monetary policy ultimately influences aggregate consumption, investment, and ouput. With some minor changes, our theory could be reinterpreted as a model of trade in an input that is reallocated across production units with different productivities. This alternative formulation would imply a more direct transmission of money shocks to the real economy. Second, our model offers a leverage-based theory of transaction velocity, a variable that is emprically relevant yet difficult to model satisfactorily in more conventional monetary models. Third, the model implications for how asset prices and their responses to monetary policy shocks depend on credit conditions and leverage could lead to fruitful empirical work.

References

[1] Afonso, Gara, and Ricardo Lagos. "Trade Dynamics in the Market for Federal Funds." Econometrica 83(1) (January 2015): 263-313.

[2] Barro, Robert J. "The Loan Market, Collateral, and Rates of Interest." *Journal of Money, Credit and Banking* 8(4) (November 1976): 439-456.

[3] Duffie, Darrell, Nicolae Gârleanu, and Lasse Heje Pedersen. "Over-the-Counter Markets." *Econometrica* 73(6) (November 2005): 1815-1847.

[4] Gali, Jordi. *Monetary Policy, Inflation, and the Business Cycle*. Princeton University Press, 2008.

[5] Kareken, John H., and Neil Wallace. "Introduction" in *Models of Monetary Economies*, Federal Reserve Bank of Minneapolis, edited by John H. Kareken and Neil Wallace, 1980.

[6] Kiyotaki, Nobuhiro, and John Moore. "Credit Cycles." *Journal of Political Economy* 105(2) (April 1997): 211-248.

[7] Kiyotaki, Nobuhiro, and John Moore. "Liquidity and Asset Prices." *International Economic Review* 46(2) (May 2005): 317-349.

[8] Kiyotaki, Nobuhiro, and Randall Wright. "On Money as a Medium of Exchange." *Journal of Political Economy* 97(4) (August 1989): 927-954.

[9] Krishnamurthy, Arvind, and Annette Vissing-Jorgensen. "The Aggregate Demand for Treasury Debt." *Journal of Political Economy* 120(2) (April 2012): 233-267.

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

[11] Lagos, Ricardo, and Shengxing Zhang. "Monetary Exchange in Over-the-Counter Markets: A Theory of Speculative Bubbles, the Fed Model, and Self-Fulfilling Liquidity Crises." NBER Working Paper 21528, September 2015.

[12] Lagos, Ricardo, and Shengxing Zhang. "Turnover Liquidity and the Transmission of Monetary Policy." New York University working paper, 2017.

[13] Lucas, Robert E. Jr. "Equilibrium in a Pure Currency Economy." *Economic Inquiry* 18(2) (April 1980): 203-220.

[14] Lucas, Robert E. Jr. and Nancy Stokey. "Optimal Fiscal and Monetary Policy in an Economy without Capital." *Journal of Monetary Economics* 12(1) (July 1983): 55-93.

[15] Rocheteau, Guillaume, Randall Wright, and Cathy Zhang. "Corporate Finance and Monetary Policy." *American Economic Review*, 2018, forthcoming.

[16] Sidrauski, Miguel. "Inflation and Economic Growth." *Journal of Political Economy* 75(6) (December 1967): 796-810.

[17] Woodford, Michael. "Doing without Money: Controlling Inflation in a Post-Monetary World." *Review of Economic Dynamics* 1(1) (January 1998): 173-219.

[18] Woodford, Michael. *Interest and Prices*. Princeton University Press, 2003.

[19] Zhu, Tao, and Neil Wallace. "Pairwise Trade and Coexistence of Money and Higher-return Assets." *Journal of Economic Theory* 133(1) (March 2007): 524-535.

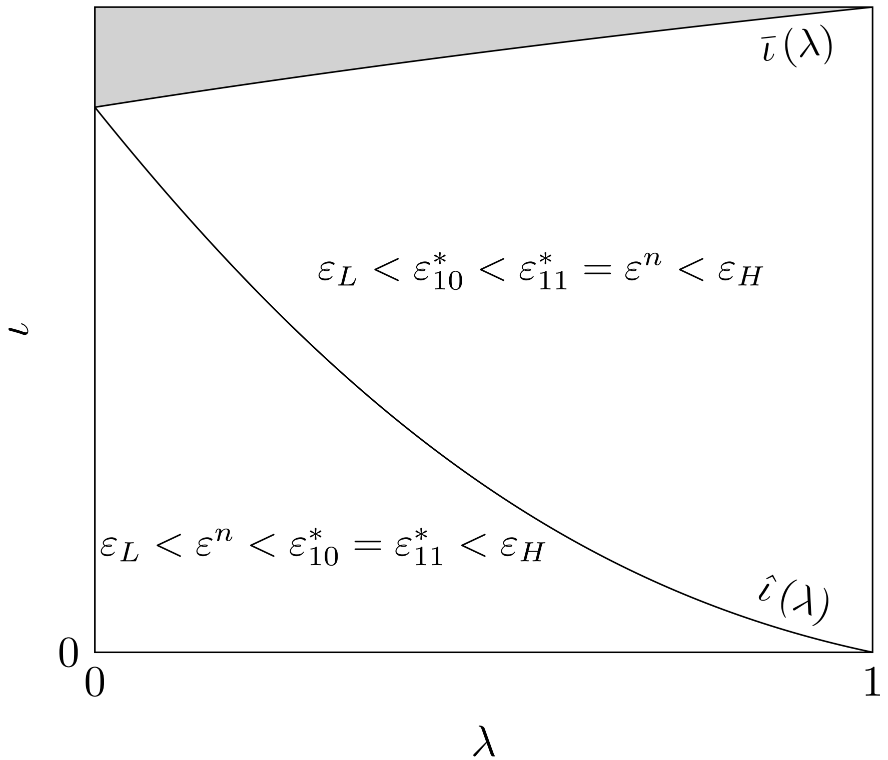


Figure 1: Existence regions for monetary equilibrium (model with general, continuous distribution of valuations, G). There is no monetary equilibrium in the shaded region. Nonmonetary equilibrium exists everywhere.

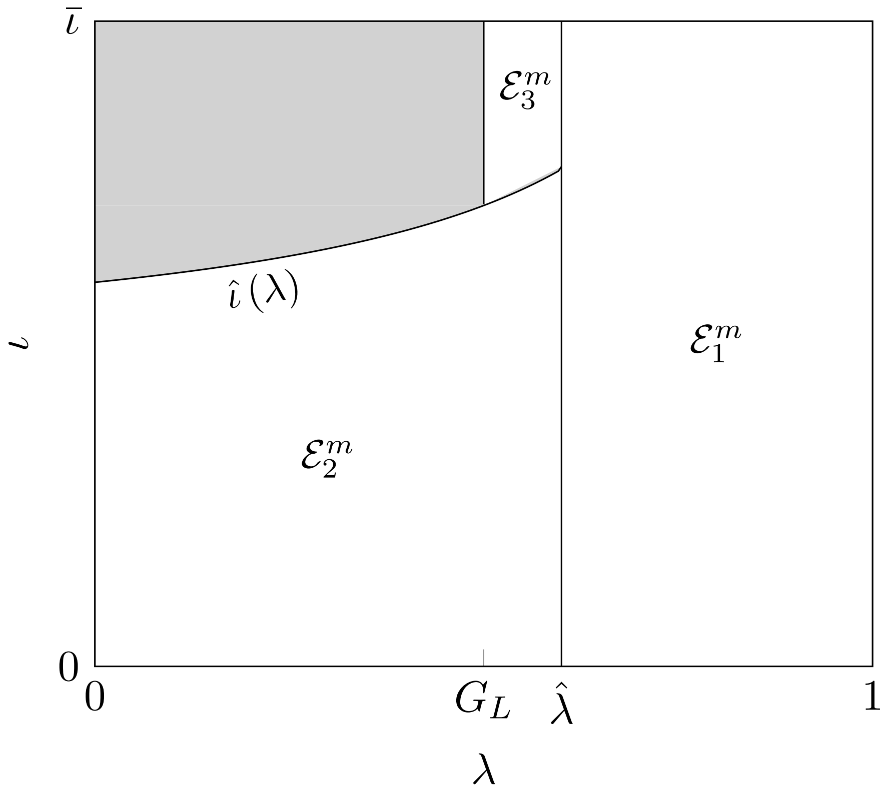


Figure 2: Existence regions for monetary equilibrium (model with two-point distribution of valuations). In region $\mathcal{E}_1^m$, the monetary equilibrium has $\varepsilon_L < \varepsilon_{10}^* < \varepsilon_{11}^* = \varepsilon_H$. In region $\mathcal{E}_2^m$, the monetary equilibrium has $\varepsilon_L < \varepsilon_{10}^* = \varepsilon_{11}^* < \varepsilon_H$. In region $\mathcal{E}_3^m$, the monetary equilibrium has $\varepsilon_L = \varepsilon_{10}^* < \varepsilon_{11}^* < \varepsilon_H$. There is no monetary equilibrium in the shaded region. Nonmonetary equilibrium exists everywhere.

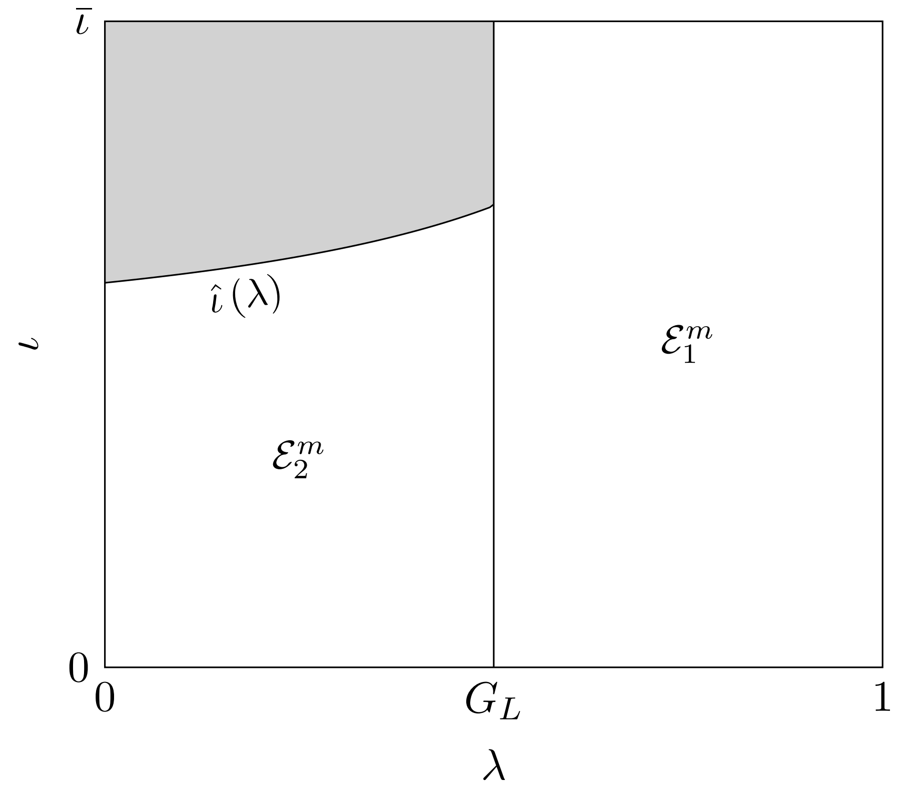


Figure 3: Existence regions for monetary equilibrium (model with two-point distribution of valuations and $\alpha_{01} = 0$). In region $\mathcal{E}_1^m$, the monetary equilibrium has $\varepsilon_L < \varepsilon_{10}^* < \varepsilon_{11}^* = \varepsilon_H$. In region $\mathcal{E}_2^m$, the monetary equilibrium has $\varepsilon_L < \varepsilon_{10}^* = \varepsilon_{11}^* < \varepsilon_H$. There is no monetary equilibrium in the shaded region. Nonmonetary equilibrium exists everywhere.

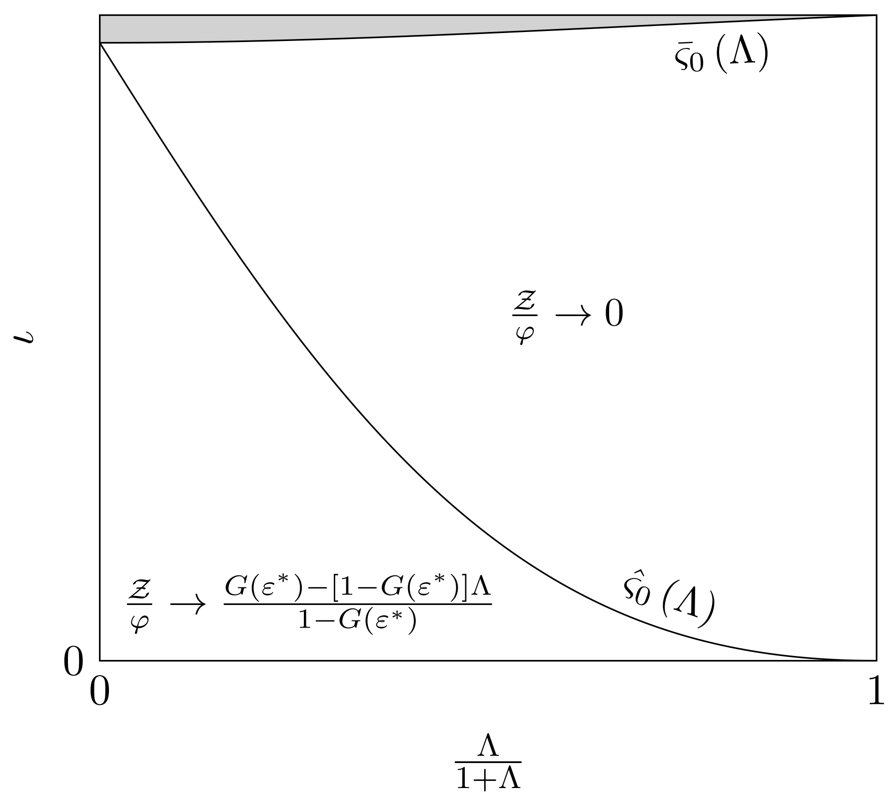


Figure 4: Existence regions for monetary equilibrium in the limit as $\alpha \to 0$ for the economy with unsecured credit (general, continuous distribution of valuations, G). There is no monetary equilibrium in the shaded region. Nonmonetary equilibrium exists everywhere. Along a monetary equilibrium, real balances $\mathcal{Z}$ converge to 0 if $\hat{\varsigma}_0(\Lambda) < \iota < \bar{\varsigma}_0(\Lambda)$ and to a positive level if $0 < \iota \leq \hat{\varsigma}_0(\Lambda)$.

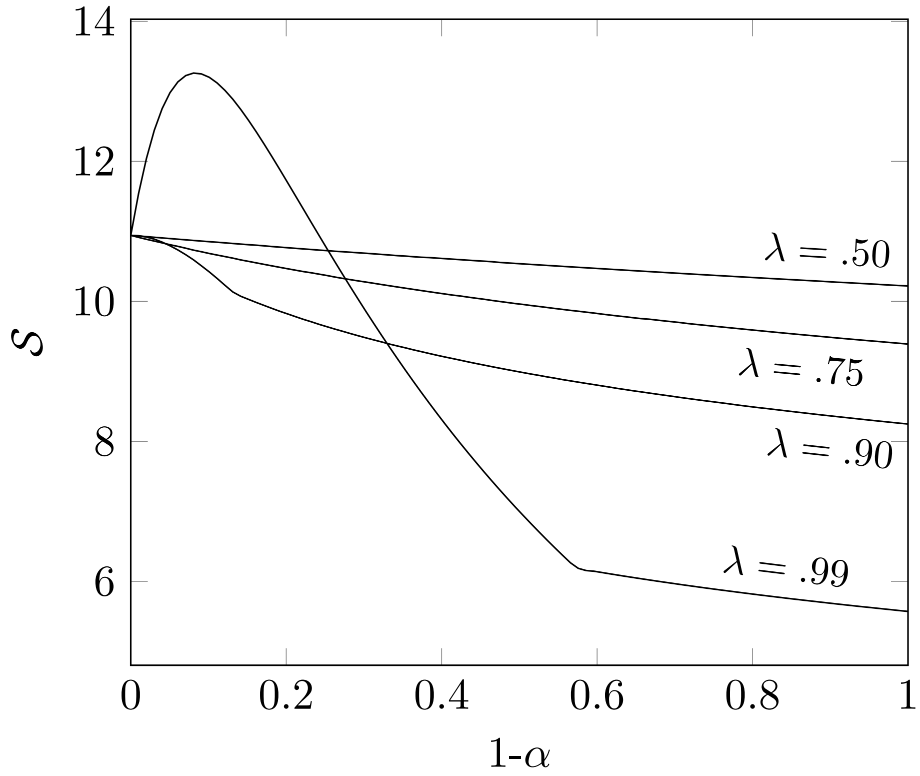


Figure 5: Semi-elasticity of the asset price with respect to the nominal policy rate for economies with different levels of leverage, λ, and access to credit, $1-\alpha$.

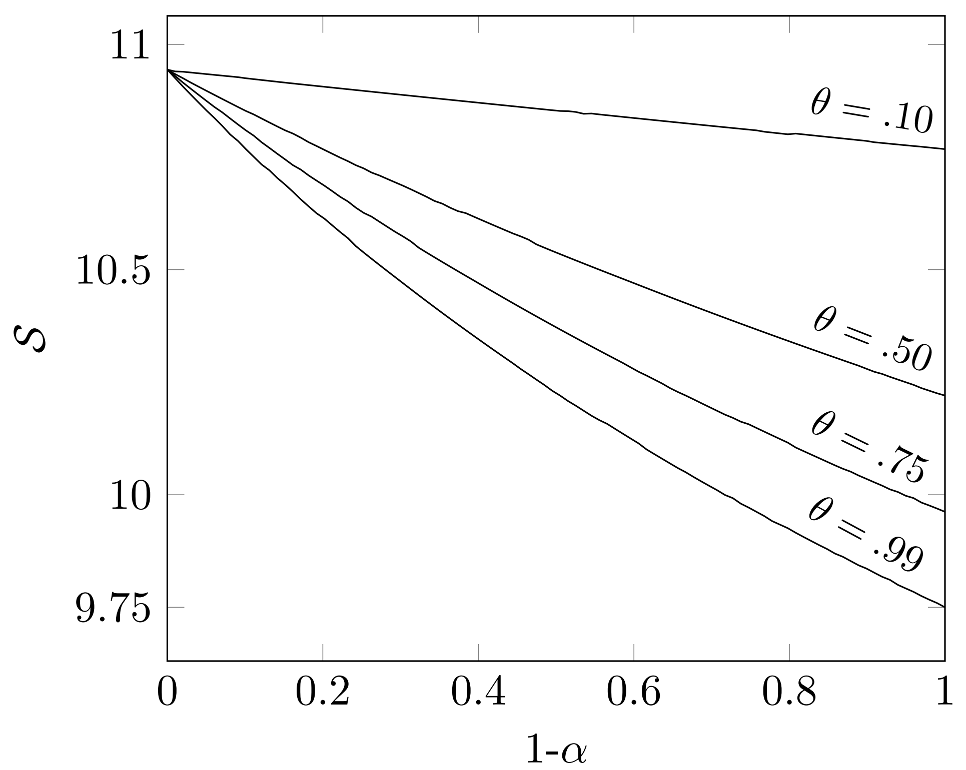


Figure 6: Semi-elasticity of the asset price with respect to the nominal policy rate for economies with different market power of brokers, $1-\theta$, and access to credit, $1-\alpha$.

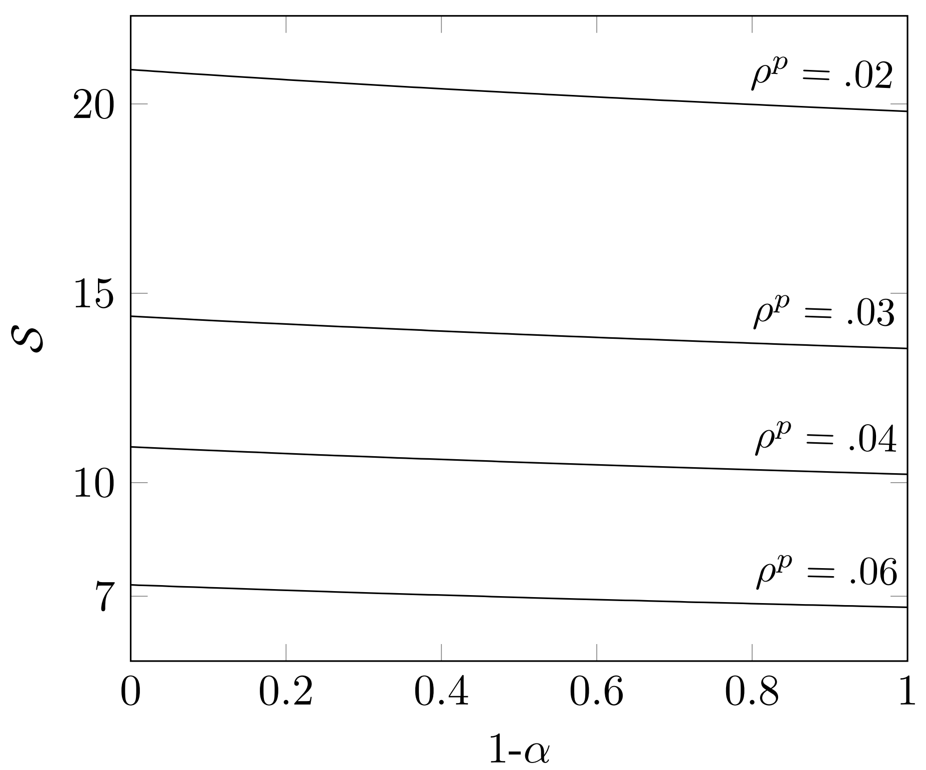


Figure 7: Semi-elasticity of the asset price with respect to the nominal policy rate for economies with different monetary regimes, ρ^p, and access to credit, $1 - \alpha$.

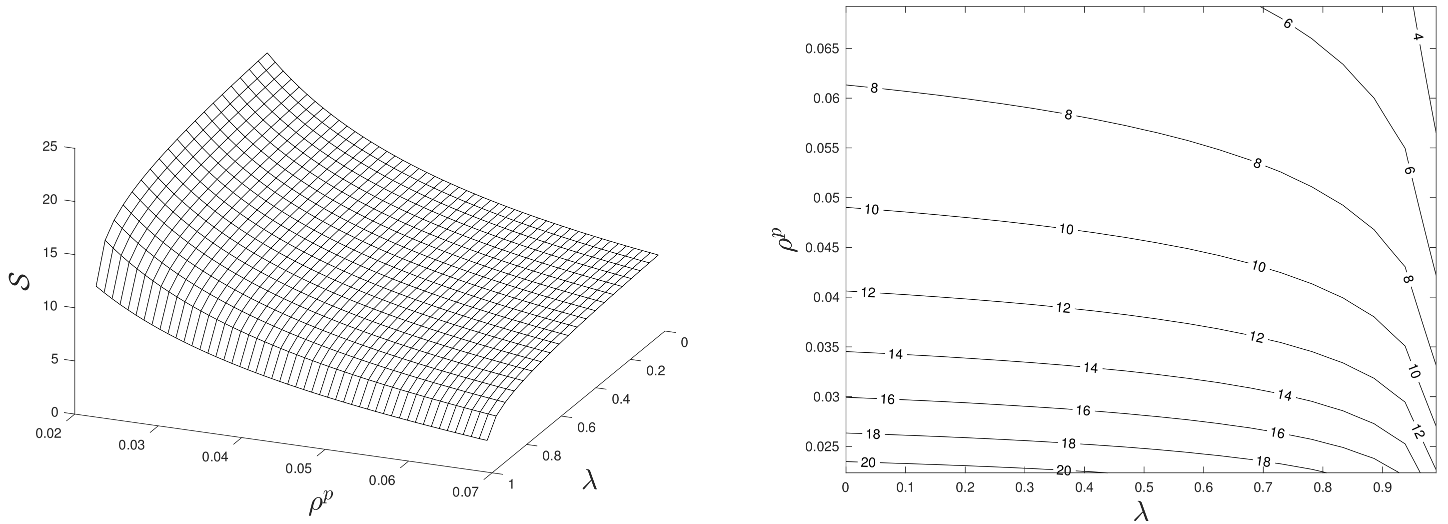
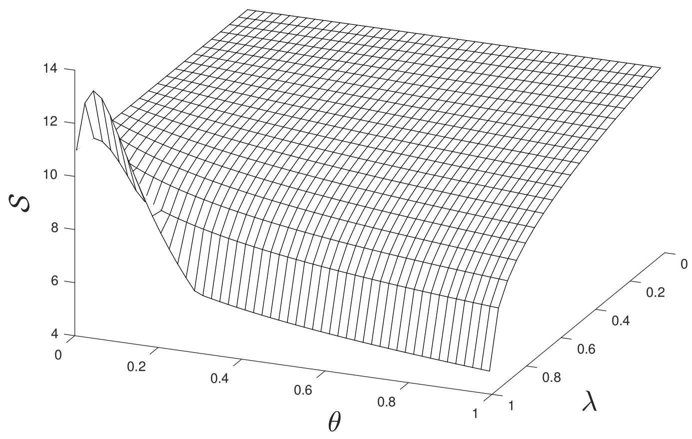


Figure 8: Semi-elasticity of the asset price with respect to the nominal policy rate as functions of λ and ρ^p in limiting economies with $\alpha \to 0$. The right panel shows the level sets for $\mathcal{S}$ corresponding to the left panel.



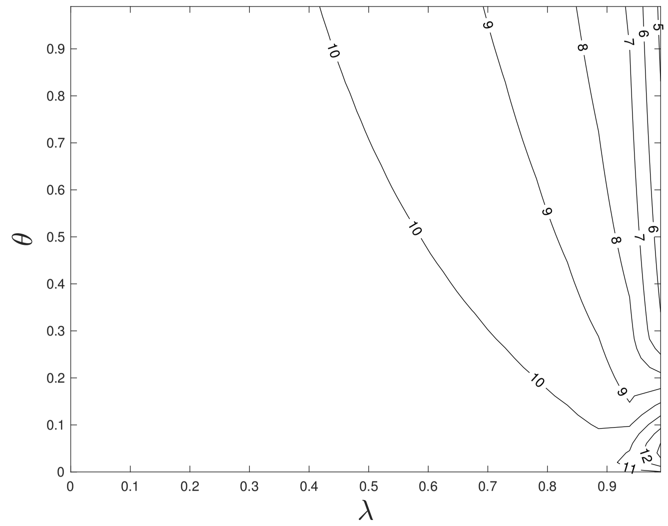


Figure 9: Semi-elasticity of the asset price with respect to the nominal policy rate as functions of λ and θ in limiting economies with $\alpha \to 0$. The right panel shows the level sets for $\mathcal{S}$ corresponding to the left panel.

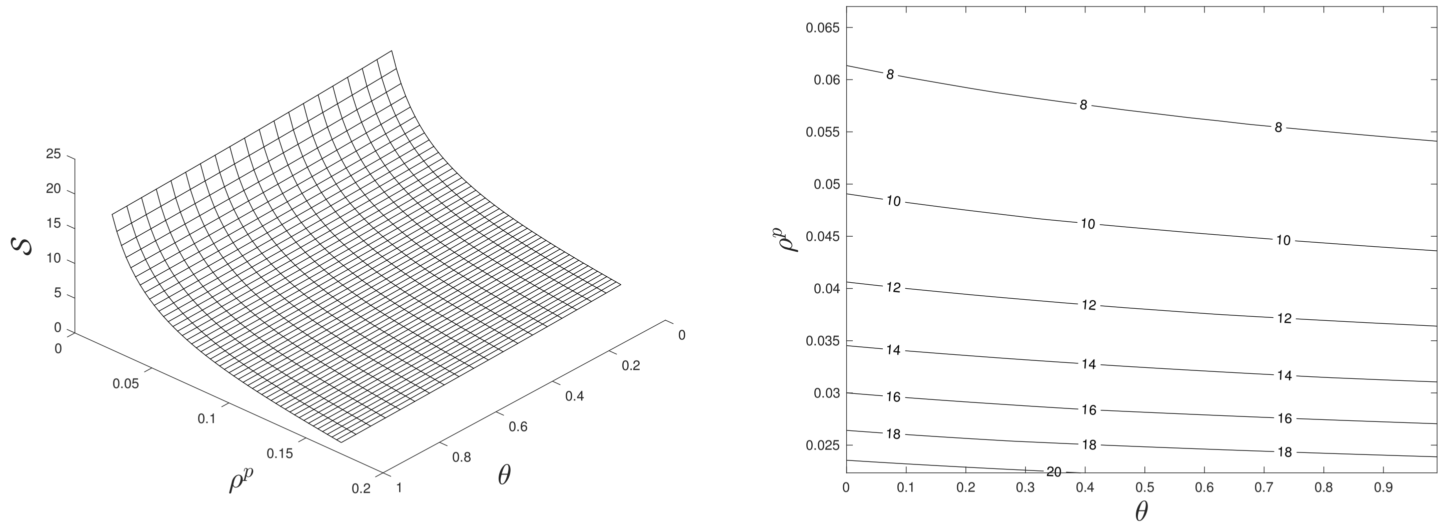


Figure 10: Semi-elasticity of the asset price with respect to the nominal policy rate as functions of θ and ρ^p in limiting economies with $\alpha \to 0$. The right panel shows the level sets for $\mathcal{S}$ corresponding to the left panel.